

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 4, 2005, 09:25:48 ; Search time 6803 Seconds  
(without alignments)  
3183.813 Million cell updates/sec

Title: US-09-851-138C-51  
Perfect score: 447  
Sequence: 1 gacggaataatttcgcaac.....aggactgcaactgttcctc 447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hhg.\*
- 3: gb\_in.\*
- 4: gb\_lm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sv.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.8	99.7	447	A50396	A50396 Sequence 51
2	445.8	99.7	447	ARI27536	ARI27536 Sequence
3	445.8	99.7	447	L39317	L39317 Hepatitis C
4	393	87.9	1584	D49749	D49749 Hepatitis C
5	393	87.9	1584	D49750	D49750 Hepatitis C
6	393	87.9	1584	D49752	D49752 Hepatitis C
7	393	87.9	1584	D63821	D63821 Hepatitis C
8	389.8	87.2	1584	D49747	D49747 Hepatitis C
9	388.2	86.8	1584	D49753	D49753 Hepatitis C
10	322.8	72.2	411	D30046	D30046 Hepatitis C
11	255.8	57.2	1505	D16614	D16614 Hepatitis C
12	252.2	56.4	1505	D16620	D16620 Hepatitis C
13	244.4	54.7	867	D16736	D16736 Hepatitis C
14	241.2	54.0	1504	D16616	D16616 Hepatitis C
15	239.4	53.6	447	L39298	L39298 Hepatitis C
16	238.2	53.3	1504	D16618	D16618 Hepatitis C
17	236.6	52.9	1504	D16612	D16612 Hepatitis C
18	236.4	52.9	867	D16737	D16737 Hepatitis C
19	236.2	52.8	541	A40613	A40613 Sequence 13

20	236.2	52.8	541	A40617	A40617 Sequence 17
21	236.2	52.8	541	BD172130	BD172130 New seque
22	236.2	52.8	541	BD172132	BD172132 New seque
23	236.2	52.8	541	AX031591	AX031591 Sequence
24	236.2	52.8	541	AX031595	AX031595 Sequence
25	236.2	52.8	541	AX031861	AX031861 Sequence
26	236.2	52.8	541	AX031865	AX031865 Sequence
27	236.2	52.8	541	AX032131	AX032131 Sequence
28	236.2	52.8	541	AX032135	AX032135 Sequence
29	236.2	52.8	541	14 HPCOREEH	D14603 Hepatitis C
30	236.2	52.8	1546	14 HPCSTRUCTC	L12355 Hepatitis C
31	234.6	52.5	541	A40619	BD172133 New seque
32	234.6	52.5	541	BD172133	BD172133 Sequence
33	234.6	52.5	541	AX031597	AX031597 Sequence
34	234.6	52.5	541	AX031867	AX031867 Sequence
35	234.6	52.5	541	AX032137	AX032137 Sequence
36	234.6	52.5	1786	14 HPCN2L1CE	D14305 Hepatitis C
37	234.6	52.5	9456	14 HPCEGS	D17763 Hepatitis C
38	233	52.1	541	A40615	A40615 Sequence 15
39	233	52.1	541	A40621	A40621 Sequence 21
40	233	52.1	541	A40625	A40625 Sequence 25
41	233	52.1	541	A40627	A40627 Sequence 27
42	233	52.1	541	BD172131	BD172131 New seque
43	233	52.1	541	BD172134	BD172134 New seque
44	233	52.1	541	BD172136	BD172136 New seque
45	233	52.1	541	BD172137	BD172137 New seque

ALIGNMENTS

RESULT 1  
A50396  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

A50396  
Sequence 51 from Patent WO9613590.  
A50396  
A50396.1 GI:2303407  
unidentified  
unclassified.  
1 (bases 1 to 447)  
Maertens, G. and Stuyver, L.  
NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS  
PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC AGENTS  
Patent: WO 9613590-A 51 09-MAY-1996;  
INNOGENETICS NV (BE)  
Other publication AU 3844095 960523.  
Location/Qualifiers  
1..447  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

DNA  
linear  
PAT 07-MAR-1997

ORIGIN	Query Match	Best Local Similarity	Mismatches	Conservative	Indels	Gaps
	99.7%	Score 445.8; DB 6;	Length 447;	100.0%;	Pred. No. 9.6e-109;	0;
				0;		
Qy	1	GACGGAATTAATTCGCAACGAGGAATTACCTGGTGGCTCTTCTCTATCTTCCTTCG	60			
Db	1	GACGGAATTAATTCGCAACGAGGAATTACCTGGTGGCTCTTCTCTATCTTCCTTCG	60			
Qy	61	GCTTTGTTCTCATGCTTGTACACCCAGCCGGGCTGGAGTACCGTAATGCTCCGGA	120			
Db	61	GCTTTGTTCTCATGCTTGTACACCCAGCCGGGCTGGAGTACCGTAATGCTCCGGA	120			
Qy	121	CTCTACATGGTAACTAACGACTGCGAGTACCGTAGTATCGTGTATGAGCCGGGATATT	180			
Db	121	CTCTACATGGTAACTAACGACTGCGAGTACCGTAGTATCGTGTATGAGCCGGGATATT	180			
Qy	181	ATCCTCCCACTTACCTGGCTGTGCTCCCTGCGTACCGTCTGGCAATACATCAAGATCTCG	240			

```
Db 181 ATCCTCCACTTACTGGCTGTGTCCTCGGTAGCTCTGGCAATACATCAAGATGCTGG 240
QY 241 ATCCCTGTGAGCCYACCGCGCCGTGAAGTGCCTTCTGGCCGCCACCGCTCTCTCCGG 300
Db 241 ATCCCTGTGAGCCYACCGCGCCGTGAAGTGCCTTCTGGCCGCCACCGCTCTCTCCGG 300
QY 301 ACCCAGTGGATATGTTGGGCGGCCACCTATGCTAGCTCTCTAGTAGAGAC 360
Db 301 ACCCAGTGGATATGTTGGGCGGCCACCTATGCTAGCTCTCTAGTAGAGAC 360
QY 361 CTTTGTGGAGCGCTATTTCTTGTGCGCGAGGGTCTCATGAGACATGCCAGCATGG 420
Db 361 CTTTGTGGAGCGCTATTTCTTGTGCGCGAGGGTCTCATGAGACATGCCAGCATGG 420
QY 421 ACTGTCAGAGTGCAGTGTTCCTATC 447
Db 421 ACTGTCAGAGTGCAGTGTTCCTATC 447

RESULT 2
AR127536
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
source
ORIGIN
Query Match 99.7%; Score 445.8; DB 6; Length 447;
Best Local Similarity 100.0%; Pred. No. 9.6e-109;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTCGCAACAGGGAATTTACTGGTGTCTTTCTCTATCTTCTCTG 60
Db 1 GACGGAATTAATTCGCAACAGGGAATTTACTGGTGTCTTTCTCTATCTTCTCTG 60
QY 61 GCTTTGTTCTCATGTTCTTACACCCAGCGGCTGGAGTACCGTAATGCTCTCCGGA 120
Db 61 GCTTTGTTCTCATGTTCTTACACCCAGCGGCTGGAGTACCGTAATGCTCTCCGGA 120
QY 121 CTCTACATGGTAACACTAGCTAGCTAGTATGCTATGAGCGCGGGATATT 180
Db 121 CTCTACATGGTAACACTAGCTAGCTAGTATGCTATGAGCGCGGGATATT 180
QY 181 ATCTTCCACTTACCTGGCTGTGTCCTCGGTAGCTCTGGCAATACATCAAGATGCTGG 240
Db 181 ATCTTCCACTTACCTGGCTGTGTCCTCGGTAGCTCTGGCAATACATCAAGATGCTGG 240
```

```
Db 421 ACTGTCAGGACTGCAACTGTTCCTATC 447

RESULT 3
HPCOREREAL
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
source
ORIGIN
Query Match 99.7%; Score 445.8; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 9.6e-109;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTCGCAACAGGGAATTTACTGGTGTCTTTCTCTATCTTCTCTG 60
Db 1 GACGGAATTAATTCGCAACAGGGAATTTACTGGTGTCTTTCTCTATCTTCTCTG 60
QY 61 GCTTTGTTCTCATGTTCTTACACCCAGCGGCTGGAGTACCGTAATGCTCTCCGGA 120
Db 61 GCTTTGTTCTCATGTTCTTACACCCAGCGGCTGGAGTACCGTAATGCTCTCCGGA 120
QY 121 CTCTACATGGTAACACTAGCTAGCTAGTATGCTATGAGCGCGGGATATT 180
Db 121 CTCTACATGGTAACACTAGCTAGCTAGTATGCTATGAGCGCGGGATATT 180
QY 181 ATCTTCCACTTACCTGGCTGTGTCCTCGGTAGCTCTGGCAATACATCAAGATGCTGG 240
Db 181 ATCTTCCACTTACCTGGCTGTGTCCTCGGTAGCTCTGGCAATACATCAAGATGCTGG 240
```

```
Db 181 ATCTCCACTTACCTGGCTGTGTCCCTCGTAGGCTCTGGCAATACATCAAGATGCTGG 240
Qy 241 ATCCCTGTGAGCCVACCGTCCGCTGAAGTGCCTCGCGCGCCACCGCCTCTCTCCGC 300
Db 241 ATCCCTGTGAGCCVACCGTCCGCTGAAGTGCCTCGCGCGCCACCGCCTCTCTCCGC 300
Qy 301 AGCCAGTGGATATGATGTGGGCGGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 301 AGCCAGTGGATATGATGTGGGCGGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Qy 361 CTTTGTGGAGCGCTATTTCTTGTGGCGAGGGGTTCTCATGGAGACATGCCAGCATTTGG 420
Db 361 CTTTGTGGAGCGCTATTTCTTGTGGCGAGGGGTTCTCATGGAGACATGCCAGCATTTGG 420
Qy 421 ACTGTCCAGGACTGCAACTGTTCCATC 447
Db 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

RESULT 4
HPCJK049A5 1584 bp RNA linear VRL 10-FEB-1999
LOCUS Hepatitis C virus isolate JK049 gene for core, env, and part of
DEFINITION E2/NS1, partial cds.
ACCESSION D49749.1 GI:1197110
VERSION core, env, and part of E2/NS1.
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (sites)
AUTHORS Tokita.H., Okamoto.H., Iizuka.H., Kishimoto.J., Tsuda.F.,
Lesmana.L.A., Miyakawa.Y. and Mayumi.M.
TITLE Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
JOURNAL 96226020
MEDLINE 8627233
PUBMED 8627233
REFERENCE 2 (bases 1 to 1584)
AUTHORS Okamoto.H.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1584)
AUTHORS Okamoto.H.
JOURNAL Direct Submission
TITLE Submitted (17-MAR-1995) Hiroaki Okamoto, Jichi Medical School,
Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi
329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
LOCATION/Qualifiers
1. .1584
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK049"
/db_xref="taxon:11103"
<1. .339
/citation=[2]
/evidence=not_experimental
340. .>1584
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA08583.1"
/db_xref="GI:1197111"
/translation="MSTLPKQRTIKRNINRRPDVKFPGGQIVGVYVLPKPKL
GVRAVKTSQSRKQIPRARTEGRNAQGYFPWPLNGCGWAGWLLSPRG
SRPSPGNDPRRSRNLGKVIDTTCGFDLMGYIPLVGAPVGGVARNALAHGVRALD
GFWGPNLPCGSFISFLALLSLTPTAGLEYNASGLYTVTNDGNSGIVYEAGD
VILHLPICPCVRLNNAKWTVPVSVRPGAAASLRTHVDMVMYGAATLCSALY
VGLDGLGALPLAGQFSWRHROHTVQDCNCSITPGHLTGRHMDMMNWNWSPMTLIV
SOUVRLPQTMFDLIGAHGWAGVAYYSMQGNWAKVFLVLCLFSGVDASTITGGVA
```

```
ORIGIN
Query Match 87.9%; Score 393; DB 14; Length 1584;
Best Local Similarity 91.9%; Pred. No. 1.5e-94;
Matches 411; Conservative 3; Mismatches 33; Indels 0; Gaps 0;
ASGAFTTISLFTSGAKQPLHLVN"
Qy 1 GAGCGAATTAATTTCCGCAACAGGGAATTTACCTGGTGGTCTCTTCTCTATCTCTCTCG 60
Db 817 GAGCGAATTAATTTCCGCAACAGGGAATTTACCTGGTGGTCTCTTCTCTATCTCTCTCG 876
Qy 61 GCTTTGTCTCATGCTTGTCTTACACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db 877 GCTTTGTCTCATGCTTGTCTTACACCCACAGCGGGCTGGAGTATCGTAATGCTCCGGA 936
Qy 121 CTCTACATGGTAACCTAACGACTGTCAGTAACGGTAGTAGTATGATGAGCGCGGGATATT 180
Db 937 CTCTACACGGTAACCAACGACTGTAGTAACGGTAGCATCGTATGAGCGCGGGATGTT 996
Qy 181 ATCTCCACTTACCTGGCTGTGTCCCTCGTAGGCTCTGGCAATACATCAAGATGCTGG 240
Db 997 ATCTCCACTTGGCTGTGTATCCCTCGTAGGCTCTTAAACAATGCATCAAAATGCTGG 1056
Qy 241 ATCCCTGTGAGCCVACCGTCCGCTGAAGTGCCTCGCGCGCCACCGCCTCTCTCCGC 300
Db 1057 ACCCTGTGAAGCCCTACCGTCCGCTGAGTGCCTCGCGCGCTACCGCCTCTCTCCGC 1116
Qy 301 AGCCAGTGGATATGATGTGGGCGGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 1117 AGCCAGTGGACATGATGTGGGAGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAT 1176
Qy 361 CTTTGTGGAGCGCTATTTCTTGTGGCGAGGGGTTCTCATGGAGACATGCCAGCATTTGG 420
Db 1177 CTTTGTGGAGCGCTATTTCTTGTGGCGAGGGGTTCTCATGGAGACATGCCAGCATTTGG 1236
Qy 421 ACTGTCCAGGACTGCAACTGTTCCATC 447
Db 1237 ACTGTCCAGGATGCAACTGTTCCATC 1263

RESULT 5
HPCJK055A6 1584 bp RNA linear VRL 10-FEB-1999
LOCUS Hepatitis C virus isolate JK055 gene for core, env, and part of
DEFINITION E2/NS1, partial cds.
ACCESSION D49750.1 GI:1197114
VERSION core, env, and part of E2/NS1.
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (sites)
AUTHORS Tokita.H., Okamoto.H., Iizuka.H., Kishimoto.J., Tsuda.F.,
Lesmana.L.A., Miyakawa.Y. and Mayumi.M.
TITLE Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
JOURNAL 96226020
MEDLINE 8627233
PUBMED 8627233
REFERENCE 2 (bases 1 to 1584)
AUTHORS Okamoto.H.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1584)
AUTHORS Okamoto.H.
JOURNAL Direct Submission
TITLE Submitted (17-MAR-1995) Hiroaki Okamoto, Jichi Medical School,
Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi
329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
LOCATION/Qualifiers
1. .1584
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK049"
/db_xref="taxon:11103"
<1. .339
/citation=[2]
/evidence=not_experimental
340. .>1584
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA08583.1"
/db_xref="GI:1197111"
/translation="MSTLPKQRTIKRNINRRPDVKFPGGQIVGVYVLPKPKL
GVRAVKTSQSRKQIPRARTEGRNAQGYFPWPLNGCGWAGWLLSPRG
SRPSPGNDPRRSRNLGKVIDTTCGFDLMGYIPLVGAPVGGVARNALAHGVRALD
GFWGPNLPCGSFISFLALLSLTPTAGLEYNASGLYTVTNDGNSGIVYEAGD
VILHLPICPCVRLNNAKWTVPVSVRPGAAASLRTHVDMVMYGAATLCSALY
VGLDGLGALPLAGQFSWRHROHTVQDCNCSITPGHLTGRHMDMMNWNWSPMTLIV
SOUVRLPQTMFDLIGAHGWAGVAYYSMQGNWAKVFLVLCLFSGVDASTITGGVA
```

```
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK055"
/db_xref="taxon:11103"
<1..339
/citation=[2]
/evidence=not_experimental
340..>1584
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA0584.1"
/db_xref="GI:1197115"
/translation="MSTLPKQKTKRNTNRRPQDVKFPFGGQIVGVVYVPRRGPKL
GVRAVHKTSERSQPRSRQPIPRARTEGRSWAOPGYWPVLYGNEGCWAGWLLSPRG
SRPSPGNDPRRSRLNKVITLTCGFADLMGYIPLVGA PVGGVARALAHGVRALD
GINFATNLPCCSFIFLLALLSCLLTPTAGLEYNASGIYVYVNDSCNSIVYEAGD
IILHLPGCVPCVRAGTSKWTPISTVAVSRPGATASLRTHVDMVMWNAATLCSALY
VGDLCGALFLVGGGFSWRHQHTVQDCNSIYPGHLTGHRAVMDMMNNSPVTMVV
SOVLRPQTIFDLVAGAHGVMAGVAYYSMQGNWAKVFLVLCFLSGVDATTITAGSA
AHSTRGFASLFTTGARQNQLVN"

ORIGIN
Query Match      87.9%; Score 393; DB 14; Length 1584;
Best Local Similarity 91.9%; Pred. No. 1.5e-94;
Matches 411; Conservative 3; Mismatches 33; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTTCTCTG 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 817 GACGGAATCAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTTCTCTG 876
QY 61 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAAATGCTTCGGA 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 877 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAAATGCTTCGGA 936
QY 121 CTCTACATGTAACACTACGACTGACGTAACGTAAGTATCGTGTATGAGCCGCGGATATT 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 937 CTCTACATGTAACACTACGACTGACGTAACGTAAGTATGTAAGTATGAGCCGCGGATATT 996
QY 181 ATCTCCACTTACCTGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 997 ATCTCCACTTACCTGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
QY 241 ATCTCCACTGAGCCYACCGTGGCGTGAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1057 ACCCTGTAAAGCCCTACCGTGGCGTGAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
QY 301 ACGCAGCTGGATATGATGGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1117 ACGCATGTGATATGATGGTGGGCGCGCCACCTATGCTCAGCTCTCTATGTAGGAGAT 1176

RESULT 6
HPVJK070A8
LOCUS
DEFINITION Hepatitis C virus isolate JK070 RNA linear VRL 10-FEB-1999
E2/NS1, partial cds.
ACCESSION D49752
VERSION D49752.1 GI:1197162
KEYWORDS core, env, and part of E2/NS1.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
```

```
Qy 361 CTTTGTGGAGCGCTATTTCTTCTGCGCAGGGGTTCTCATGGAGACATGCCAGCATTTGG 420
Db 1177 CTTTGTGGAGCGCTTTCTTCTGTTGGCAGGGGTTCTCATGGAGACATGCCAGCATTTG 1236

Qy 421 ACTGTCCAGGACTGCAACTGTTTCCATC 447
Db 1237 ACTGTTTCAGGAGTGCACAACTGTTTCGATC 1263

RESULT 7
LOCUS HPCJK049E1 9450 bp RNA linear VRL 13-FEB-1999
DEFINITION Hepatitis C virus isolate.JK049 genomic RNA for polyprotein,
complete genome.
ACCESSION D63821
VERSION D63821.1 GI:1183032
KEYWORDS polyprotein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
1 (sites)
REFERENCE Tokita.H., Okamoto,H., Iizuka.H., Kishimoto,J., Tsuda.F.,
AUTHORS Lesmana,L.A., Miyakawa,Y. and Mayumi.M.
TITLE Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
JOURNAL J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
MEDLINE 96226020
PUBMED 8627233
REFERENCE 2 (bases 1 to 9450)
AUTHORS Okamoto,H.
JOURNAL Unpublished
-REFERENCE 3 (bases 1 to 9450)
AUTHORS Okamoto,H.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-1995) Hiroaki Okamoto, Jichi Medical School,
Immunology Division; Minamikawachi-machi, kawachi-gun, Tochigi
329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
FEATURES
Location/Qualifiers
1..9450
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK049"
/db_xref="taxon:11103"
1..9450
/feature="genomic RNA for polyprotein precursor, complete
genome"
/citation=[2]
/evidence=not_experimental
1..339
/citation=[2]
/evidence=not_experimental
340..9399
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="polyprotein"
/protein_id="BAA09890.1"
/db_xref="GI:1183033"
/translation="MSTLPKFORITKRNINRRPQDKPFGGQIVGVVIVPRRPGKL
GVRAVTKRSRSQSRQPIPRARTEGRSQAQPGYWPPLYGNEGCGWAGLLSPRG
SRSPGPNDRPRSRNLGKVIDLTLCGADLMGYIPLVGAQVGVGARNALAHGVRLED
GINPATLPGCSFSIFLLALLSCLLTPTAGLEYRNAGSLVTVTNDSCNSIYVBAQD
VILHPGCI PCVRLANNAKSWTPVSVAPGGAATSLRTHVDMVGAATLCSALY
VGDLLPGALFLVQGGFSWRHQHWTVQDCNSIYPGHLTCHRWMDMNMWSPMTLIV
SOVLKPLPTMFDLIGAHGVNAGVAYYSMQGNWAKVFLVLCFLPSGVDASTITTGVA
ASGAFITLSLFSRGAQPLHLVNTGNSWHINRTALNCNDSLNTFAGLLYYHKFNSS
GCVRMASCSPLDRPAQWGLPGPANI SGPSKEPKYCHWYAPRCDTVPAQSVCCGPV
CFTPSVVGATDKRGATPTTYTNGENSDVFLLESAR PTEPWFGCTWNGSGYVKTG
APPHICVGGREGKSNLVCPDTCERKHPDATNRCGAGPMLT PRLVDVYPRLWHYP
CTVNTYIFKVRMFVGGLEHFRFNAACNWRGERCNLEDRDRSEMYFLLSHSTTEQALLPC
```

```
SPVPIPALSTGLIHLHQNVQVLYGLSSGLVGHAIKWEFVILLFLLADARVCVVL
MMMLISQAREALENDLVLNALSATAGTHGIMNSLVAFCAVMHVRGIFPIAVISVIGL
WPLLILMLPYRAYATWTGTDTSITAGVFLSLFALPTLSPMYKHWIARLIWYIIVGL
RCEALQIWPPPLARGARDGIIILAGLFYPALVFDITKLLAIIIGLPYILQASLVRV
PFVFRHAVVRLCILVRNITGKVKYQVMVLLALARGFNTLYDLHLGSPMTDWAAGLKL
AVAVEPVI FSPMEVKVI TWGADTTAGCIIICGLPVSAIRGKVLGPPADDYRSMGWL
LAPITAHAAQOTRGLPTGTVTSITGPDKNITVGEIOVLSSTOTFLGTSGVGVMTVYH
GAGSRTLAGNKRPAALQMTNVDDQLVGNPSPGAKSLVPTCTGSDADLYLITRDADVL
ARRRGDSTASLLSPRLACLKSGSGGPIIMCPSGHVAGIFRAAVCTRGVAKALQFIPVE
SLSAQTRSPFSFSDNSTPAPVQTFQVGYLHATGTGSESTKVPASVAAQYTVLVNPS
VAATLQGRFMSHAYIGIDNVNRTGRTTITAKLTYSYKFLDAGGSGGVAITGEEI
DECHAQDATSLIGITVDQOAEATAGARLWATATPGSITVPHSNIENDEAVLTGEBE
PFYGRAI PLGVITKGRHLI FCHSKKCDLAKOLTSLVNAVAFVYRGLDVSIVIPQGD
VVCATDALITGYTGDPSIDCNVAVEQYVDFSLDPTFSEIETHVTPQDARSQRGG
RTGSGTGYTSPGERSGDFSVLCECTDAGCATGATFETPTSETTIDRAYLISLTPG
LPVCDHLEFWEVGTGLTHIDHAFSLQTKQGLNFPYLYTAYQATCARAAALPPSWD
ETWKLIRLPTLHGPTPLLYRLAGVONEICTTHPVTYIATCMADLEVAATSAWLL
GGVMAALTAYCLSVGSVVI VGLVGLGKPALVPDKVELYQOYDEMECSRAAPYTEQA
OGIAQOFKEKVLGILLQADOKAADI KPIATPVMOKLETPWSKHMNPVSGIOYLAGLS
TLPGNPALASLMATASVTSPLTNTQILLNMGWASNLAPPASTAFVYSLAGA
AVSGTGLKVLDDILAGYGAGVAGALVAFKIMGHVPSTEDMVNLLPALLSPALVVG
VICAILRRHVGPGEVAVOMNRLIAFASRGNHVPATHYVSPESDAAKYATLLSLTV
TQLLRLHQWINEDEYTPCDGNWLYDIWNWCTVLADFKMLGAKILPMPGIPFLSC
OKYRGTWGQGVWSTPCGALLSGHVKNQTMRLVGRWCANTWHTGTFPIINGYTTGP
STPAPSYASRALMRVADS YVEVKVGDHFVYCTTDDGLKCPQVPLPERFETLDG
VLRHYAVPCRPLLADDDVTFTVGLNSYVIGSQLPCEPDVAVVTSMLODPSHITVET
AKRLDRGSPPLSSASSASQSAFSRKATCTTHGRHPDAELITALLMRQWGSNITR
VESESKVILDSFELPRACDEDELSAAECFKPKPYPPALPIWARPDPNPLPVEPM
KDPDYVPTVHGCAIPPOKLPVPPRRKRTIVLSESTVSKASLASAEKSPQPCTSA
BDESTGVGTOSGLTGPVOLDSDNESHSMPLEGEQDPLDSSGSMWTSVGESE
QSVVCCSMSYNTGALITPCAABEEKLPISPLNSLARHNLVYTSRSRAQOKV
TFDRLOVDHNTYITLKEIKELASGVKAEELSVEACRLVPSHSARSFGYGAKEVRS
LSSKAINHINSWEDLEDNTPTPTTMAXNEVFAVAPHKGGKRPALIVYDPLGVR
ICEKRALYDVIVKLPISAIMSAYGQFYSQKORVELLKMWNKSKTPLGFSYOTRCPDS
LPTQEDIRVEESIYQACDLKDEARRVITSLTERLYCGGPMKQHCQYRRCRASG
TVTEQNTVTCVLKAKAATKAAGIKDPSFLCGDLDLVIAESAGIIDEKSAALAPTEA
MTRYSAPPDPPQPTVDELITSCSSNVSAHDGKBYVYLTRDEPTEPLAANETA
RHTPVNSMIGNIIMAPTIVWRVMIMTHFFSILQAOEQLEKALDEMGVANGVSVTPLD
LPALIERLHGLSAFSLHSYSPVELNRVAGRLKLGIPPLRAWHRBARAKLISQGG
KAKICGLYFNWAVRTAKLTPLQAGILLDSRNFVTGAGGNDIYHSVRSRARSRLLL
GLLLLTGVGVIFLLPAR"
9400..9450
/note="with 17nt of T-stretch"
/citation=[2]
/evidence=not_experimental

ORIGIN
3'UTR
Query Match 87.9%; Score 393; DB 14; Length 9450;
Best Local Similarity 91.9%; Pred. No. 1.7e-94;
Matches 411; Conservative 3; Mismatches 33; Indels 0; Gaps 0;

Qy 1 GACGGAATTAATTTTCGCAACAGGGGAATTTACCTGGTGTCTCTTTCTATCTCTCTCTG 60
Db 817 GACGGAATCAATTTTCGCAACAGGGGAATTTACCTGGTGTCTCTTTCTATCTCTCTCTG 876

Qy 61 GCTTTGTTCTCATGCTTACACCCACAGCGGGCTGGAGTACCGTAATGCTCTCCGA 120
Db 877 GCTTTGCTCTCATGCTTACACCCACAGCGGGCTGGAGTATCGTAATGCTCTCCGA 936

Qy 121 CTCTACATGTAACCTAACGACTGACGTAGTATCGTGTATGAGCGCGGGGATATT 180
Db 937 CTCTACACGGTAACCAACGACTGTAGTAAACGCTAGCATCGTGTATGAGCGCGGGGATGTT 996

Qy 181 ATCTCTCACTTACCTGGCTGTGTCCCTTCGCTACGCTCTGGCAATACATCAAGATGCTGG 240
Db 997 ATCTCTCACTTGGCTGGCTGTATCCCTGGCTGGCTTTAAACAATGATCAAAATGCTGG 1056

Qy 241 ATCCCTGTGAGCCCVACCGTCCGCGTGAAGTCGCCCTCGGCCGCCGCCCGCTCTCTCCGC 300
Db 1057 ACCCTGTGAAGCCCTTACCGTCCGCGTGAAGTCGCCCTCGGCCGCCGCCCGCTCTCTCCGC 1116

Qy 301 AGCGAGTGGATATGATCGTGGGCGGCCACCTTATCTCAGCTCTCTACGTAGAGAC 360
```

```

Db      1117  ACGCAGTGGACATGATGGTGGAGCGGCCACCCCTAAGCTCAGCTCTCTACGTAGGAGAT 1176
QY      361  CTTTGTGAGGCTATTTCTTGTGGGAGGGGTTCTCATGAGACATGCGCAGCATTTGG 420
Db      1177  CTTTGTGAGGCTATTTCTTGTGGGAGGGGTTCTCATGAGACATGCGCAGCATTTGG 1236
QY      421  ACTGTCAGGAGTCAACTGTTCCATC 447
Db      1237  ACTGTCCAGGATTGCAACTGTTCCATC 1263

RESULT 8
LOCUS   HPCJK030A3
DEFINITION Hepatitis C virus isolate JK030 gene for core, env, and part of
ACCESSION D49747
VERSION D49747.1 GI:1197102
KEYWORDS core, env, and part of E2/NS1.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
REFERENCE 1 (sites)
AUTHORS Tokita,H., Okamoto,H., Iizuka,H., Kishimoto,J., Tsuda,F.,
Lesmana,L.A., Miyakawa,Y. and Mayumi,M.
TITLE Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
JOURNAL J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
MEDLINE 96226020
PUBMED 8627233
REFERENCE 2 (bases 1 to 1584)
AUTHORS Okamoto,H.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1584)
AUTHORS Okamoto,H.
JOURNAL Direct Submission
TITLE Submitted (17-MAR-1995) Hiroaki Okamoto, Jichi Medical School,
Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi
329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
LOCATION/Qualifiers
FEATURES
source 1. .1584
/mol_type="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK030"
/db_xref="taxon:11103"
<1. .339
/citation=[2]
/evidence=not_experimental
340. .>1584
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA08581.1"
/db_xref="GI:1197103"
/translation="MSLTPKQRITKNTNRPPQDKVPGGGQIVGGVYVLRPGPKL
GYRAVKTSESRQRNRQIPPRARTGRSQAQPGYPWLYNGEGCGWAGLLSPRG
SRPWSGNDPRRRNLGKVIDTLTCGFADLMGYIPLVGAPVGVARALAHGRLLED
GINFATNLPCFSFIPLALLSCLLTPAAGLEYRNASGLYMTVNDCSNIIIVEAGD
IILHLPQVPCVRSNGTSKWTSVPTVAVSHPGAATASLRTHVDMVGNATLCSALY
VGDLCLGFLVGGQFSKRRQHWTVQDCNSIYPGHLTGRHMAWMMWNSPATLVV
SOVLRLPQTLIDLVIGHWGMAGVAYISMQGNWAKRFLVLCLFSGVDASTRISGSA
AHNTYGLSSLFSSGPKNQIQLIN"

Query Match 87.2%; Score 389.8; DB 14; Length 1584;
Best Local Similarity 91.5%; Pred. No. 1.1e-93;
Matches 409; Conservative 3; Mismatches 35; Indels 0; Gaps 0;

QY      1 GACCGAATTAATTCGCAACAGGAATTACCTGGTGTCTCTCTATCTTCTCTG 60

```

```

Db      817  GACCGAATCAATTTTGCACACAGGAATTACCTGGTGTCTCTCTATCTTCTCTG 876
QY      61  GTTTTGTCTCATGCTTGTACACCCACAGCCGGGCTGGAGTACCGTAAGTCCCTCCGGA 120
Db      877  GCTTTGTCTCATGCTTGTACACCCACAGCCGGGTTGGAGTACCGTAATGCCCTCCGGA 936
QY      121  CTCTACATGGTAACTAACGACTGCGAGTAACGGTAGTATCGTGTATGAGCCGGGATATT 180
Db      937  CTCTACATGGTAAACCAACGACTGTAGCAACAGTAGTATCGTGTATGAGCCGGGATATT 996
QY      181  ATCTCTCCACTTACCTGGCTGTGCTCCCTGCGCTGCGCTCTGCGCAATACATCAAGATGCTGG 240
Db      997  ATCTCTCCACTTACCTGGCTGTGCTCCCTGCGCTGCGCTCTGCGCAACATCAAAATGCTGG 1056
QY      241  ATCCCTGTGAGCCCYACCGTGCCTGCAAGTCGCGCCCTGCGCGCCACGCGCTCTCTCCGC 300
Db      1057  ACCTCTGTGAAGCCCTACTGTGCGCGTGAAGTACACCTGTGCTGCGCACCGCTCTCTACGC 1116
QY      301  ACGCAGCTGGATATGATGTTGGGCGGCCACCCATGCTCAGCTCTCTACGTAGGAGAC 360
Db      1117  ACACAGCTGGATATGATGTTGGGCGGCCACCCATGCTCAGCTCTCTACGTAGGAGAT 1176
QY      361  CTTTGTGAGCGCTATTTCTTGTGGGACGGGTTCTCATGGAGACATCGCCAGCATTTGG 420
Db      1177  CTTTGTGAGCGCTATTTCTTGTGGACAGGGTTCTCATGGAAACGTGCCAGCATTTGG 1236
QY      421  ACTGTCCAGGACTGCAACTGTTCCATC 447
Db      1237  ACTGTCCAGGATTGCAACTGTTCCATC 1263

RESULT 9
LOCUS   HPCJK072A9
DEFINITION Hepatitis C virus isolate JK072 gene for core, env, and part of
ACCESSION D49753
VERSION D49753.1 GI:1197124
KEYWORDS core, env, and part of E2/NS1.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
REFERENCE 1 (sites)
AUTHORS Tokita,H., Okamoto,H., Iizuka,H., Kishimoto,J., Tsuda,F.,
Lesmana,L.A., Miyakawa,Y. and Mayumi,M.
TITLE Hepatitis C virus variants from Jakarta, Indonesia classifiable
into novel genotypes in the second (2e and 2f), tenth (10a) and
eleventh (11a) genetic groups
JOURNAL J. Gen. Virol. 77 (Pt 2), 293-301 (1996)
MEDLINE 96226020
PUBMED 8627233
REFERENCE 2 (bases 1 to 1584)
AUTHORS Okamoto,H.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1584)
AUTHORS Okamoto,H.
JOURNAL Direct Submission
TITLE Submitted (17-MAR-1995) Hiroaki Okamoto, Jichi Medical School,
Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi
329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
LOCATION/Qualifiers
FEATURES
source 1. .1584
/mol_type="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="JK072"
/db_xref="taxon:11103"
<1. .339
/citation=[2]
/evidence=not_experimental
340. .>1584
/evidence=not_experimental

```

```
/citation=[2]
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA08587.1"
/db_xref="GI:1197125"
/translation="MSTLPKPKRKTRNTNRPPQVDPKPGGQIVGVVYVLPKPKL
GVRAVRKTSRSPRSRQPIPRARRTGSRWAOPGYPBLYGNEGCGWGLLSPRG
SRPSGNDPRRSRNLKVIDTLTCGFADLMGYIPLAGAPVGVVARALAGVRALED
GINFATNGPSCFSIFLLALLSLTPTAGLEYRNSGLYVNTDSCNSISIVYEATD
IILHPGVCPCVRSNTSRWISISPTVAVSHPGATASLRTHVDMVMYGAATLCSALY
IGDLCGAFVLVQGSFMRHRLQWTVQECNCSYIPGHLTCHRMAMDMNMWSPAVTMV
SOVLRLPOTILDVIGAHGWAGVAYYSMQGNWAKVFLVLCFLPSGVDAHTTTTGGVA
ARTTSGLTALTFTGPKQHQLIN"
```

## ORIGIN

```
Query Match      86.8%; Score 388.2; DB 14; Length 1584;
Best Local Similarity 91.3%; Pred. No. 3e-93;
Matches 408; Conservative 3; Mismatches 36; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTCCGACACAGGGAATTTACCTGGTGTCTTTCTCTATCTCTCTCTG 60
DB 1 GACGGAATTAATTTCCGACACAGGGAATTTACCTGGTGTCTTTCTCTATCTCTCTCTG 876
QY 61 GCTTTGTTCTCATGCTTGTACACCCACAGCGGGCTGGAGTACCGTAATGCCCTCCGGA 120
DB 877 GCTTTACTCTCATGCTTGTACACCCACAGCTGGGCTAGATACCGTAACAGCTCCGGA 936
QY 121 CTCTACATGGTAACACTGACAGTAAACGGTAGTATCGTGTATGAGCGCGGGGATATT 180
DB 937 CTCTACATGGTAACCAACGACTGTAGTAACAGTAGTATCGTGTATGAGCGCGGATATC 996
QY 181 ATCTCTCACTTACCTGGCTGTGTCCTCGGTACGCTCTGGAATACATCAAGATGCTGG 240
DB 997 ATCTCTCACTTACCTGGCTGTGTCCTCGGTACGCTCTGGAATACATCAAGATGCTGG 1056
QY 241 ATCTCTGAGCCVACCGTGCCTGGAAGTCCGCTGGCGCGCCACCGCTCTCTCCGC 300
DB 1057 ATCTCTAAGCCCTTACTGTGCGCGGTAGTATCTGCGGCTGCGCCACCGCTCTCTCCGC 1116
QY 301 ACGCAGCTGGATATGATGGTGGCGCGCCACCTTATGCTCAGCTCTCTACGTAGGAGAC 360
DB 1117 ACGCAGCTGGATATGATGGTGGCGCGCCACCTTATGCTCAGCTCTCTACATAGAGAT 1176
QY 361 CTTTGTGAGCGCTATTTCTTGTGGGAGGCGGTTCTCATGAGACATCGCCAGCATGG 420
DB 1177 CTTTGTGAGCGATATTTCTTGTGGGAGGCGGTTCTCATGAGACATCGCCAGCTTGG 1236
QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447
DB 1237 ACTGTTCAGGAGTGCAACTGTTTCGATC 1263
```

## RESULT 10

```
HPCP3          HPCP3          411 bp      RNA      linear      VRL 07-FEB-1999
LOCUS          Hepatitis C virus (individual isolate fd-3/93) gene for polyprotein
DEFINITION     precursor, partial cds (core protein (carboxy terminus) and E1
                envelope protein (amino terminus half)).
ACCESSION      D30046.1 GI:485798
VERSION        D30046.1 GI:485798
KEYWORDS       E1 envelope protein; core protein.
SOURCE         Hepatitis C virus
ORGANISM       Hepatitis C virus
                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                Hepacivirus.
REFERENCE      1 (sites)
AUTHORS        Hotta,H., Handajani,R., Lusida,M.I., Soemarto,W., Doi,H.,
                Miyajima,H. and Homma,M.
TITLE          Subtype analysis of hepatitis C virus in Indonesia on the basis of
                NS5b region sequences
JOURNAL        J. Clin. Microbiol. 32 (12), 3049-3051 (1994)
MEDLINE        95189942
```

## PUBMED

2 (bases 1 to 411)

Hotta,H.

Unpublished

3 (bases 1 to 411)

Hotta,H.

Direct Submission

Submitted (28-APR-1994)

Medicine, Department of Microbiology;

Kobe, Hyogo 650, Japan (Tel:078-341-7451(ex.3301),

Fax:078-351-6347)

Submitted (28-Apr-1994)

to DDBJ by:

Hak Hotta

Kobe University School of Medicine

Department of Microbiology

7-5-1 Kusunoki-cho, Chuo-ku

Kobe, Hyogo 650

Japan

Phone: 078-341-7451 x3301

Fax: 078-351-6347

Location/Qualifiers

1. .411

/organism="Hepatitis C virus"

/mol\_type="genomic RNA"

/isolate="Td-3/93"

/db\_xref="taxon:11103"

&lt;1. .&gt;411

/note="The carboxy terminus of the core protein and the

amino terminal half of the E1 envelope protein of

hepatitis C virus"

/codon\_start=1

/product="polyprotein precursor"

/protein\_id="BAA06282.1"

/db\_xref="GI:485799"

/translation="LSCILLTPTAGLEYRNASGLYIVTNDSCNSISIVYEAQDILHMPG

CDPCVRSNTSRWTSPTVAVSHPGATASLRTHVDMVMYGAATLCSALYVGDLCGA

LFLVQGSFMRHRLQWTVQECNCSYIPGHLTCHRM"

## ORIGIN

```
Query Match      72.2%; Score 322.8; DB 14; Length 411;
Best Local Similarity 90.0%; Pred. No. 9.9e-76;
Matches 342; Conservative 3; Mismatches 35; Indels 0; Gaps 0;

QY 68 TCTCATGCTTCTTACACCCACAGCGGGCTGGAGTACGTAATGCCCTCCGACTCTACA 127
DB 2 TCTCATGCTTCTTACACCCACAGCGGGCTGGAGTATCGTAATGCCCTCCGACTCTACA 61
QY 128 TGGTAACCTAACGACTGACGTAACGTAAGTATCGTGTATGAGCGCGGGATATTATCCTCC 187
DB 62 TAGTAACCAACGACTGACGTAACGTAAGTATCGTGTATGAGCGCGGAGATATTATCCTCC 121
QY 188 ACTTACCTGGCTGTGTCCCTCGCTTGGCAATACATCAAGATGCTGGATCCCTG 247
DB 122 ACATGCTTGGCTGTGACCCCTCGCTTGGCAACACATCAAGGTGCTGGACCCCTG 181
QY 248 TGAGCCCVACGCTGCGCGTGAAGTGGCCCTGCGCGCCACCGCTCTCTCCGACGACG 307
DB 182 TAAGCTCTACTGTGCTGTGGGTGCGCTGGCGCTGTCTACCGCTCTCTCCGACACATG 241
QY 308 TGGATATGATGTTGGGCGGCGCCACCTTATGCTCAGCTCTCTACGTAGGAGACCTTTG 367
DB 242 TGGATATGATGTTGGGCGGCGCCACCTTTCTCTAGCTCTCTCTACGTAGGAGATCTT 301
QY 368 GAGCGCTATTTCTTGTGGGCGAGGGTCTCTCATGGAGACATCGCCAGCATTTGGACTG 427
DB 302 GAGCGCTTTTCTTGTGGGCGAGGGTCTCTCATGGAGACATCGCCAGCATTTGGACTG 361
QY 428 AGGACTGCAACTGTTCCATC 447
DB 362 AGGATTGCAACTGTTCCATC 381
```

## RESULT 11



```

HPCNE125C
LOCUS      HPCNE125C      1505 bp      RNA      linear      VRL 10-JUN-1999
DEFINITION Hepatitis C virus RNA for core, env, and part of E2/NS1 polyprotein.
ACCESSION  D16614
VERSION    D16614.1  GI:475866
KEYWORDS   E2/NS1; core protein; envelope protein.
SOURCE     Hepatitis C virus
ORGANISM   Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE  1 (sites)
AUTHORS   Shrestha,S.M., Tsuda,F., Okamoto,H., Tokita,H., Horikita,M.,
            Tanaka,T., Miyakawa,Y. and Mayumi,M.
TITLE     Hepatitis B virus subtypes and hepatitis C virus genotypes in
            patients with chronic liver disease in Nepal
JOURNAL   Hepatology 19 (4), 805-809 (1994)
MEDLINE   94186155
PUBMED    8138250
REFERENCE  2 (sites)
AUTHORS   Tokita,H., Shrestha,S.M., Okamoto,H., Sakamoto,M., Horikita,M.,
            Iizuka,H., Shrestha,S., Miyakawa,Y. and Mayumi,M.
TITLE     Hepatitis C virus variants from Nepal with novel genotypes and
            their classification into the third major group
JOURNAL   J. Gen. Virol. 75 (Pt 4), 931-936 (1994)
MEDLINE   94201770
PUBMED    8151307
COMMENT   Submitted (06-JUL-1993) to DDBJ by:
            Hiroaki Okamoto
            Immunology Division
            Jichi Medical School
            Kawachi-gun
            Tochigi-ken 329-04
            Japan
            Phone: 0285-44-2111 x3334
            Fax: 0285-44-1557.
FEATURES   source
            Location/Qualifiers
            1..1505
               /organism="Hepatitis C virus"
               /mol_type="genomic RNA"
               /isolate="NE125"
               /db_xref="taxon:11103"
               <1..260
               /evidence=not_experimental
            261..>1505
               /codon_start=1
               /evidence=not_experimental
               /product="core, env, and part of E2/NS1"
               /protein_id="BAA04036.1"
               /db_xref="GI:475867"
               /translation="MSTPKPQRPKRNTPRPQDVKPPGGQIVGGVYVLPKRGPRL
               GVRVRSKRSRQPRGRQPIPKARRSGRSWAQPGYFWPLYNEGCGWAGLLSPRG
               SRPSGPNDRRRNRNUGKVIDTLTCGFDALMGYIPLVGAPVGGVARALAHGVRLED
               GINFGALPGCSFSIFLLALFSLVSPAGLEYRNTSGLYILTNDCSNSIIVYEADN
               VILHMPGVCPTREGNQRCTVPTPTVAVKHPGAVTASIRRHVDLMVGAATLCSALY
               VGLDCAFLVGOAFTFRARQHTVQLCNSLYPHGITHGHMAWDMNWSPTAALVV
               AHLRIPTVDFMIVGAHWGMAGLAFMQGNWAKVGIIVLLFSGVDADTCTTGGA
               ARNARGFASLSSGNRQQLQV"
ORIGIN
Query Match      57.2%; Score 255.8; DB 14; Length 1505;
Best Local Similarity 73.6%; Pred. No. 1e-57;
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY      1  GACGGAATTAATTCGCAACAGGAAATTACCTGGTGTCTTTCTCTATCTTCCTCTG 60
Db      738 GACGGAATTAATTCGCAACAGGAAATTGCGCGGTGCTCTTTCTATCTTCCTCTT 797

QY      61  GCTTGTGTTCTCATGTTGCTTACACCCACAGCCGGGTGGAGTACCGTAAATGCTCGGA 120
Db      798 GCTCTCTTCTCATGTTGTTTCCCGCGCAGCGGGGTAGATACAGGAACAAGTCGCGC 857

QY      121 CTCTACATGGTAACACGATGACGATGACGATGATGATGATGATGATGATGATGATGAT 180

```

```

Db      858 CTATACATATTACCAACAGACTGTCTTAACAGCAGCATCGTGTATGAGGCTGATAATGTC 917
QY      181 ATCTTCCCACTTACCTGGCTGTGTCCCTGCGTAGCTCTTGGCAATACATCAAGATGCTGG 240
Db      918 ATCTTGCACATGCCCGGCTGTGTGCTCTGCACTCGCGAGGGTAACCAAGTCAAGGTGCTGG 977
QY      241 ATCCCTGTGAGCCCVACCGTCGCGTGAAGTCGCGCTCGCGCCCTGCGCCAGCCGCTCTTCGCG 300
Db      978 ACGCAGTAACACCGACAGTGGCTGTCAAACATCTGCGCGAGTCACCGCATCAATCCGC 1037
QY      301 ACGCAGTGGATATGATGTTGGGCGCGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db      1038 AGGCATGTGATTTGATGTTGGTGGTGACGCCACGCTGTGTTTCAGCACTCTATGTTGGAGAT 1097
QY      361 CTTTGTGAGCGCTATTCTTCTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATGG 420
Db      1098 TTGTGCGGGCTGTTTCTTGTGGGCAAGCGTTCACTTTTCAGAGTCGCGCAACATTAT 1157
QY      421 ACTGTCCAGGACTGCAACTGTTC 443
Db      1158 ACGTCCAGTTGTGCAATTGCTC 1180

RESULT 12
HPCNE274I
LOCUS      HPCNE274I      1505 bp      RNA      linear      VRL 10-JUN-1999
DEFINITION Hepatitis C virus RNA for core, env, and part of E2/NS1 polyprotein.
ACCESSION  D16620
VERSION    D16620.1  GI:475878
KEYWORDS   E2/NS1; core protein; envelope protein.
SOURCE     Hepatitis C virus
ORGANISM   Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE  1 (sites)
AUTHORS   Shrestha,S.M., Tsuda,F., Okamoto,H., Tokita,H., Horikita,M.,
            Tanaka,T., Miyakawa,Y. and Mayumi,M.
TITLE     Hepatitis B virus subtypes and hepatitis C virus genotypes in
            patients with chronic liver disease in Nepal
JOURNAL   Hepatology 19 (4), 805-809 (1994)
MEDLINE   94186155
PUBMED    8138250
REFERENCE  2 (sites)
AUTHORS   Tokita,H., Shrestha,S.M., Okamoto,H., Sakamoto,M., Horikita,M.,
            Iizuka,H., Shrestha,S., Miyakawa,Y. and Mayumi,M.
TITLE     Hepatitis C virus variants from Nepal with novel genotypes and
            their classification into the third major group
JOURNAL   J. Gen. Virol. 75 (Pt 4), 931-936 (1994)
MEDLINE   94201770
PUBMED    8151307
COMMENT   Submitted (06-JUL-1993) to DDBJ by:
            Hiroaki Okamoto
            Immunology Division
            Jichi Medical School
            Kawachi-gun
            Tochigi-ken 329-04
            Japan
            Phone: 0285-44-2111 x3334
            Fax: 0285-44-1557.
FEATURES   source
            Location/Qualifiers
            1..1505
               /organism="Hepatitis C virus"
               /mol_type="genomic RNA"
               /isolate="NE274"
               /db_xref="taxon:11103"
               <1..260
               /evidence=not_experimental
            261..>1505
               /codon_start=1
               /evidence=not_experimental
               /product="core, env, and part of E2/NS1"
               /protein_id="BAA04042.1"
               /db_xref="GI:475879"
            5' UTR
            CDS

```



```

/translation="MSTLPKPQKTKRNTIRPQDKVPPGGQIVGVYLLPRGRPLP
SVFVAKTISPSRNRPQIPKARSSEGRSQAQYQWPLYNKGEGCWAGWLLSPRG
GSRWNTNPRRRSRNRLVQITLTCGADLMGYPVLGVPGVQVAGALAHGVRALD
GINFATNPGCSFSITPLAALISCLVCPAAGYEVNVSGLVLTNDCSNGISVYEAP
VILHGPCVPCVQSGNSSCWTPTAVKAYAGATIASISHDVLDLGAATLCSALY
VQCMGKAVFLPQKQTPVQNTVQTCNSLYPGHISGRHMDMMNMMNSPVLGV
AHLVRPQTIFITAGAHGYNAGLAIYYSMQTNNAKVPILIVNFSGYDARTHTTTGGV
HTTSLTITFTRGPSGNQLIN"

```

ORIGIN	Query Match	Best Local Similarity	Best Match	Score	252.2	DB 14	Length	1505	;	
	Matches	323	Conservative	1	Mismatches	119	Indels	0	Gaps	0
Qy	1	GACGGAATTAATTTCGCAACAGGGAATTTACCTGGTTCCTTTCTCTATCTCTCTCTCTG	60							
Db	738	GACGGAATAAATTTGCAACAGGGAATTTACCCGGTTGCTTTCTCTATCTCTCTCTG	797							
Qy	61	GCITTTGTTCTCATGCTTTGCTTTACACCCACAGCCGGCTGGAGTACCGTAAATGCTCCGGA	120							
Db	798	GCITTTGCTCTCTGCTTTGGTCTGCTCTGCTGCAGGGAATTAATACCCGGAATGTCTGGC	857							
Qy	121	CTCTACATGTAACATAACGACTGCAGTAACCGGTAGTATCGTGTATGAGCGCGGGATATT	180							
Db	858	CTCTACGTGCTCACCAACGACTGCTCTAACGGCAGTATCGTATGAGGCCCTCTGATGC	917							
Qy	181	ATCCTCCACTTACCTGCTGTGTCCCTCGGTACGCTCTGGCCAATACATCAAGATGCTGG	240							
Db	918	ATCTTGACATTCGCAGGTGTGTGCCCTCGGTTCAATCAGGCAACCTCTCGCAATGCTGG	977							
Qy	241	ATCCCTGTGAGCCCYACCGTCCGGTGAAGTCCGCCCTGCCCGGCCACCGCCTCTCTCCGC	300							
Db	978	ATTCCGGTGGCCACCAACAGTGGGGCGTTAAGTACGCTGGCGCGACCACTGATCGATCCGC	1037							
Qy	301	ACGCACGTGGATATGATGTGGGCGGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC	360							
Db	1038	AGTCATGTGGATCTGCTGTGGGAGCTGCTACGTTGTGCTCCGCGCTGATGTTGGCGAT	1097							
Qy	361	CTTTGTGGACGCTATTCTTTGTGCGAGCGGTTCTCATGGAGACATCCCGACGATTGG	420							
Db	1098	ATGTGTGGACCGCTCTCTTTGGTGGACAGGCTTTCACCTTCAGACCTCGTCAGACAC	1157							
Qy	421	ACTGTCCAGACTGCAACTGTTTC	443							
Db	1158	ACGGTGCAGACCTGCAATTGCTC	1180							

```

RESULT 13
LOCUS      HPCHCV048                      linear      VRL 15-FEB-2003
DEFINITION Hepatitis C virus DNA, clone:BA-1.
ACCESSION  D16736
VERSION    D16736.1  GI:506254
KEYWORDS
SOURCE
ORGANISM   Hepatitis C virus
            Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE  1
AUTHORS   Ohno,T.
TITLE     Hepatitis C virus
JOURNAL   Thesis (1993) The University of Tokyo
REFERENCE  2  (bases 1 to 867)
AUTHORS   Ohno,T.
TITLE     Direct Submission
JOURNAL   Submitted (08-JUL-1993) Tomoyoshi Ohno, Nagoya City University
            Medical School, Second Department of Internal Medicine; 1-1
            Kawasumi, Mizuho, Nagoya, Aichi 467, Japan
            (Tel:052-851-5511(ex.8748,2265), Fax:052-852-0849)
FEATURES
            Location/Qualifiers
            1..867
            /organism="Hepatitis C virus"
            /mol type="genomic RNA"

```

```
/db_xref="taxon:11103"
/clone="BA-1"
```

Query Match	54.7%; Score 244..4; DB 14; Length 867;
Best Local Similarity	71.7%; Pred. No. 1.le-54;
Matches 320; Conservative	0; Mismatches 126; Indels 0; Gaps 0;
Qy	1 GACGGAAATTAATTTGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTCTCTCTG 60
Db	365 GACGGAGTGAATATGCAACAGGGAATCTTCCTGGTGTCTTTCTATCTCTCTCTC 424
Qy	61 GCCTTTGTTCTCATGCTTGCTTTACACCAAGACGGGGCTGGAGTACCGTATAGCTCCGGA 120
Db	425 GCTCTCTTCTCTGCTGCTGACTTCCCGCGCATCTGCTCTGGAGTACAGGAACGGTCCGGC 484
Qy	121 CTCTACATGGTAACTAACGACTCAGTAAACGGTAGTATCGTGTATAGGCGCGGGATATT 180
Db	485 CTATACATGCTTACCAACGACTGCTCTAAACAGAGCATCGTGTACGAGGCTGACAGGTG 544
Qy	181 ATCTCTGACTTACTGTGCTGTGTCCTGGTGTGCTGGGCAATACATCAAGATGCTGG 240
Db	545 ATCTGTGCATTACCCGGATGTGTGCCATGCATCAAACTGGCAACCAAAACATCGTGTGG 604
Qy	241 ATCCCTGTGAGCCCYACCGTTCGGCGTGAAGTGCCTGCCCTGCCCGCACCGGCTCTCTCCGC 300
Db	605 ACACCAAGTGTACCGACTGTGGCTGTCAAGTACCTGGTGGACCAACCGCATCGATCCGC 664
Qy	301 ACSCACGTGGATATGATGTGGGCGGCGCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db	665 AATCATGTGGATATGTTGGTGGCGGGCGGCACCTCTGTCTCGGCGCTATACGTCTGGGGAC 724
Qy	361 CTCTTGTGGAGCGCTATTTCTTGTGNGGCGGGGTCTCATGGAGACATCCCAAGCATTTG 420
Db	725 CTCTCGGGAGCCGTGTTTCTTGTGGAGCAAGCGTTTCACTTTCAGGCTCTGTGCACACAG 784
Qy	421 ACTGTCCAGACTGCAACTGTTCCAT 446
Db	785 ACCGTGCAGACGTGCAACTGCTCAAT 810

RESULT 14	HPCNE137E	1504 bp	RNA	linear	VRL 10-JUN-1999			
LOCUS	HPCNE137E	Hepatitis C virus RNA for core, env, and part of E2/NS1 polyprotein.						
DEFINITION	D16616							
ACCESSION	D16616.1	GI:475870						
VERSION	E2/NS1	core protein; envelope protein.						
KEYWORDS	Hepatitis C virus							
SOURCE	Hepatitis C virus							
ORGANISM	Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.							
REFERENCE	1 (sites)							
AUTHORS	Shrestha,S.M., Tsuda,F., Okamoto,H., Tokita,H., Horikita,M., Tanaka,T., Miyakawa,Y. and Mayumi,M.							
TITLE	Hepatitis B virus subtypes and hepatitis C virus genotypes in patients with chronic liver disease in Nepal							
JOURNAL	Hepatology 19 (4), 805-809 (1994)							
MEDLINE	94186155							
PUBMED	8138250							
REFERENCE	2 (sites)							
AUTHORS	Tokita,H., Shrestha,S.M., Okamoto,H., Sakamoto,M., Horikita,M., Iizuka,H., Shrestha,S., Miyakawa,Y. and Mayumi,M.							
TITLE	Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group							
JOURNAL	J. Gen. Virol. 75 (Pt 4), 931-936 (1994)							
MEDLINE	94201770							
PUBMED	8151307							
COMMENT	Submitted (06-JUL-1993) to DDBJ by: Hiroaki Okamoto							
	Immunology Division							
	Jichi Medical School							
	Kawachi-gun							

```
Tochigi-ken 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557.
Location/Qualifiers
1. .1504
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="NE137"
/db_xref="taxon:11103"
<1. .259
/evidence=not_experimental
260. >1504
/codon_start=1
/evidence=not_experimental
/product="core, env, and part of E2/NS1"
/protein_id="BAA04038.1"
/db_xref="GI:475871"
/translation="MSTLPKPOQTKEHTIYRQRPPQVNFPGGQIVGGVVLPRRGPRL
GVPRAVKTSSQPRGRQIPKAREGRSQAQGPWPLFNGECGWAGWLLSPRG
SRFSWGNDFRRSRNLGVITLTCGFADLMGYIPLIGAPVGGVARALAHGVRALD
GVNATGNLPGCSFSLFLALFSLCTCPASGLEVRTNSGLVYLINDCNOSIVYADE
VILHPCGVPCVATANTSCWTPVSTVACVHLGATASIRSHVDMVGAATLCSALY
VGLCCGAVLGVQAFVSPRRHATVQTCNSIVPGHISGHRWMDMMWNSPAGLAV
SHMLRPLQTFEFLIAGAHGVMAGLAYFSMQGNWAKVGLVILMFSGVDAETHVTGGTV
ASGASLITRFSNPKQLQLVN"
ORIGIN
Query Match 54.0%; Score 241.2; DB 14; Length 1504;
Best Local Similarity 71.3%; Pred. No. 8.5e-54;
Matches 318; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 1 GACGGAAATAATTTCGACACAGGAATTTACCTGGTGGCTCTTTCTCTATCTTCCTTCTG 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
737 GACGGAGTAGAATGACACACAGGAATCTTCTGGTGGCTCTTTCTCTATCTTCCTTCTT 796
QY 61 GCTTTGTTCTCATGCTTGCTTACACCCACAGCGGGCTGGAGTACCGTAATGCCCTCCGGA 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
797 GCTCTCTCTCTGCTTGACTTGGCCGGGCTCTGGTCTAGAGTACAGGAACACGCTCTGGC 856
QY 121 CTCTACATGGTAATCAACGATCGAGTAACCGTAGTATCGTGTATGAGCGCGGGGATATT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 CTCTACGCTGCTACCAACGACGTCTTAACAGAGCATCGTGTATGAGCGCGAGAGGTG 916
QY 181 ATCTCTCCACTTACCTGGCTGTGTCCTCCGCTACGCTCTGCAATACATCAAGATGCTGG 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
917 ATCTTGCACTTACCCGGATGTGTCCTCGCTGCACTGCACTGCCAACCAACATCGTCTGG 976
QY 241 ATCCCTGTGAGCCYVACCCTGCGCGTGAAGTTCGCGCGCACCGCCTCTCTCCGC 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
977 ACCCCAGTGTACCCAGCGTAGCTGTCAAGCACCTTGGTGGACACCGGTCCATCCGC 1036
QY 301 ACCGACGTGGATATGATGGTGGRCGCCCAACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1037 AGTCATGTGGATATGTGGTGGCGCGCAGCCACGTTGTGTTCAAGCGCTATACGTCGGGGAC 1096
QY 361 CTTTGTGAGCGCTATTCTTGTGTGGCAGGGGTTCTCAAGGACATATCCGACATGG 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1097 CTTTGGGGGCGGTGTTCTAGTGGCAACGCGTTCACTTCAGCCCTCGCGCGCATGG 446
QY 421 ACTGTCCAGGACTGCAACTGTTCCAT 446
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1157 ACTGTACAGACATGCAACTGCTCGAT 1182
RESULT 15
HPCOREBELS
LOCUS Hepatitis C virus type 3a clone NL26 precursor protein gene,
DEFINITION partial cds.
ACCESSION L39298
VERSION L39298.1 GI:845459
KEYWORDS
```

Db 361 ATGTGTGGGGCGTCTTCTTGTGGCACAAGCCTTCACATTGAGGCTGTGCGCCATCNA 420

Qy 421 ACTGTCCAGGACTGCAACTGTTC 443

Db 421 ACGGTCCAGACTTGTAACTGCTC 443

Search completed: March 4, 2005, 13:44:48  
Job time : 6807 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2005, 11:18:03 ; Search time 697 Seconds  
(without alignments)  
1049.377 Million cell updates/sec

Title: US-09-851-138C-51  
Perfect score: 447  
Sequence: 1 gacgaatttaattgcaac.....aggactgcaactgttccatc 447

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.8	99.7	447	3	US-08-836-075A-51
2	236.2	52.8	541	4	US-09-878-281A-13
3	236.2	52.8	541	4	US-09-878-281A-17
4	234.6	52.5	541	4	US-09-878-281A-19
5	233	52.1	541	4	US-09-878-281A-15
6	233	52.1	541	4	US-09-878-281A-21
7	233	52.1	541	4	US-09-878-281A-25
8	233	52.1	541	4	US-09-878-281A-27
9	233	52.1	630	3	US-08-612-973-29
10	233	52.1	630	3	US-08-927-597-29
11	231.4	51.8	541	4	US-09-878-281A-23
12	217.4	48.6	579	4	US-09-878-281A-177
13	215.8	48.3	574	4	US-09-878-281A-120
14	215.8	48.3	579	4	US-09-878-281A-175
15	215.8	48.3	579	4	US-09-878-281A-181
16	209.8	46.9	447	3	US-08-836-075A-39
17	209	46.8	447	3	US-08-836-075A-31
18	208.6	46.7	447	3	US-08-836-075A-37
19	206.6	46.2	579	4	US-09-878-281A-173
20	206.2	46.1	579	4	US-09-878-281A-187
21	205	45.9	579	4	US-09-878-281A-171
22	204.6	45.8	574	4	US-09-878-281A-118
23	204.2	45.7	2116	3	US-08-191-160-21
24	203	45.4	579	4	US-09-878-281A-185
25	201.4	45.1	579	4	US-09-878-281A-155
26	201.4	45.1	579	4	US-09-878-281A-195
27	201	45.0	447	3	US-08-836-075A-29

28 199.4 44.6 742 1 US-08-081-072-18 Sequence 18, Appl  
29 199.4 44.6 742 1 US-08-449-093A-18 Sequence 18, Appl  
30 199.4 44.6 932 1 US-08-081-072-15 Sequence 15, Appl  
31 199.4 44.6 932 1 US-08-449-093A-15 Sequence 15, Appl  
32 198.2 44.3 447 3 US-08-836-075A-35 Sequence 35, Appl  
33 197.8 44.3 447 3 US-08-836-075A-33 Sequence 33, Appl  
34 196.2 43.9 501 2 US-08-483-695-30 Sequence 30, Appl  
35 196.2 43.9 501 2 US-07-965-285-30 Sequence 30, Appl  
36 196.2 43.9 501 2 US-08-487-231-30 Sequence 30, Appl  
37 196.2 43.9 501 2 US-09-201-912-30 Sequence 30, Appl  
38 196.2 43.9 633 3 US-08-612-973-7 Sequence 7, Appl  
39 196.2 43.9 633 3 US-08-927-597-7 Sequence 7, Appl  
40 196.2 43.9 795 3 US-08-612-973-5 Sequence 5, Appl  
41 196.2 43.9 795 3 US-08-927-597-5 Sequence 5, Appl  
42 196.2 43.9 2082 3 US-08-612-973-47 Sequence 47, Appl  
43 196.2 43.9 2082 3 US-08-927-597-47 Sequence 47, Appl  
44 196.2 43.9 2433 3 US-08-612-973-49 Sequence 49, Appl  
45 196.2 43.9 2433 3 US-08-927-597-49 Sequence 49, Appl

## ALIGNMENTS

RESULT 1  
US-08-836-075A-51  
; Sequence 51, Application US/08836075A  
; Patent No. 6180768  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: STUYVER, LIEVEN  
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
; TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
; TITLE OF INVENTION: AGENTS  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD, WHITE & DURKEE  
; STREET: P.O. BOX 4433  
; CITY: HOUSTON  
; STATE: TEXAS  
; COUNTRY: USA  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Microsoft Word 6.0 / ASCII text output  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,075A  
; FILING DATE: 21 Apr 1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/04155  
; FILING DATE: 23 Oct 1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 94870166.9  
; FILING DATE: 21 Oct 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95870076.7  
; FILING DATE: 28 Jun 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAMMERER, PATRICIA A.  
; REGISTRATION NUMBER: 29,775  
; REFERENCE/DOCKET NUMBER: INNS:004  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-836-075A-51

Query Match 99.7%; Score 445.8; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.9e-124;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGGTCTTCTCTATCTCTCTCTG 60  
Db |||||  
QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGGTCTTCTCTATCTCTCTCTG 60  
Db |||||

QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120  
Db |||||

QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120  
Db |||||

QY 121 CTCTACATGTAATAACGACTGAGTAACGCTAGTATCGTGTATGAGCGCGGGGATATT 180  
Db |||||

QY 121 CTCTACATGTAATAACGACTGAGTAACGCTAGTATCGTGTATGAGCGCGGGGATATT 180  
Db |||||

QY 181 ATCCTCCACTTACCTGGCTGTGTCCTGCTAGCTCTGGCAATACATCAAGATGCTGG 240  
Db |||||

QY 181 ATCCTCCACTTACCTGGCTGTGTCCTGCTAGCTCTGGCAATACATCAAGATGCTGG 240  
Db |||||

QY 241 ATCCCTGTGAGCCYACCGTCCGCTGAAAGTCCGCTGGCCGCCACCCGCTCTCTCCGC 300  
Db |||||

QY 241 ATCCCTGTGAGCCYACCGTCCGCTGAAAGTCCGCTGGCCGCCACCCGCTCTCTCCGC 300  
Db |||||

QY 301 ACCCAGTGGATATGATGGTGGGCGGCCACCCCTATGCTCTCTACGTAGGAGAC 360  
Db |||||

QY 301 ACCCAGTGGATATGATGGTGGGCGGCCACCCCTATGCTCTCTACGTAGGAGAC 360  
Db |||||

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATTTGG 420  
Db |||||

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATTTGG 420  
Db |||||

QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
Db |||||

QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
Db |||||

RESULT 2  
US-09-878-281A-13  
; Sequence 13, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Innogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph  
; FILE REFERENCE: 35  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 541  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-878-281A-13

Query Match 52.8%; Score 236.2; DB 4; Length 541;  
Best Local Similarity 70.7%; Pred. No. 2.9e-61;  
Matches 313; Conservative 1; Mismatches 129; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGGTCTTCTCTATCTCTCTCTG 60  
Db |||||

QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120  
Db |||||

QY 122 GCTCTGTTCTTCTTAAATCCATCCAGCAGCTAGTCTAGAGTGGGAAACACGTCTGGC 181  
Db |||||

QY 121 CTCTACATGTAATAACGACTGAGTAACGCTAGTATCGTGTATGAGCGCGGGGATATT 180  
Db |||||

QY 182 CTCTATGCTTACCAACGACTGTTTCCAAATAGCAGTATTTGTGTATGAGCGCGGATGACGTT 241  
Db |||||

QY 181 ATCCTCCACTTACCTGGCTGTGTCCTGCTAGCTCTGGCAATACATCAAGATGCTGG 240  
Db |||||

Db 242 ATTCTGCACACACCCGGCTGTGTACTTGTGTTCAGGACGGTAATACATCTCGGTGCTGG 301  
QY 241 ATCCCTGTGAGCCYACCGTCCGCTGAAAGTGGCCCTGGCGGCCACCGCTCTCTCCGC 300  
Db |||||

Db 302 ACCCAGTGGACACCTTACAGTGGCAGTACGTCGGAGCAACACCGCTTCGATACGC 361  
QY 301 AGCAGCTGGGATATGATGGTGGGCGGCCACCCCTATGCTCTCAGCTCTCTACGTAGGAGAC 360  
Db |||||

Db 362 AGGCATGTAGACATGTTGTGGCGGCCACAGATGCTCTGTCTCTCTACGTGGGTGAT 421  
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATTTGG 420  
Db |||||

Db 422 ATGTGTGGGCGCTCTTCTCTGTGGGACAAGCCTTCAGTTTCAGACCTCGTCCCATCAA 481  
QY 421 ACTGTCCAGGACTGCAACTGTTTC 443  
Db |||||

Db 482 ACGTCCAGACCTGTAACTGCTC 504  
QY 481 ACTGTCCAGGACTGCAACTGTTTC 443  
Db |||||

RESULT 3  
US-09-878-281A-17  
; Sequence 17, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Innogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph  
; FILE REFERENCE: 35  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 541  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-878-281A-17

Query Match 52.8%; Score 236.2; DB 4; Length 541;  
Best Local Similarity 70.7%; Pred. No. 2.9e-61;  
Matches 313; Conservative 1; Mismatches 129; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGGTCTTCTCTATCTCTCTCTG 60  
Db |||||

QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120  
Db |||||

QY 122 GCTCTGTTCTTCTTAAATCCATCCAGCAGCTAGTCTAGAGTGGGAAACACGTCTGGC 181  
Db |||||

QY 121 CTCTACATGTAATAACGACTGAGTAACGCTAGTATCGTGTATGAGCGCGGGGATATT 180  
Db |||||

QY 182 CTCTACGTCTTACCAACGACTGTTTCCAAATAGCAGTATTTGTGTATGAGCGCGGATGACGTT 241  
Db |||||

QY 181 ATCCTCCACTTACCTGGCTGTGTCCTGCTAGCTCTGGCAATACATCAAGATGCTGG 240  
Db |||||

QY 242 ATTCTGCACACACCCGGCTGTGTACTTGTTCAGGACGGTAATACATCTCGGTGCTGG 301  
Db |||||

QY 241 ATCCCTGTGAGCCYACCGTCCGCTGAAAGTGGCCCTGGCGGCCACCGCTCTCTCCGC 300  
Db |||||

QY 302 ACCCAGTGGACACCTTACAGTGGCAGTACGTCGGAGCAACACCGCTTCGATACGC 361  
Db |||||

QY 301 AGCAGCTGGGATATGATGGTGGGCGGCCACCCCTATGCTCTCAGCTCTCTACGTAGGAGAC 360  
Db |||||

QY 362 AGGCATGTAGACATATTTGTGGCGGCCACAGATGCTCTGTCTCTCTACGTGGGTGAT 421  
Db |||||

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATTTGG 420  
Db |||||

QY 422 ATGTGTGGGCGCTCTTCTCTGTGGGACAAGCCTTCAGTTTCAGACCTCGTCCCATCAA 481  
Db |||||

QY 421 ACTGTCCAGGACTGCAACTGTTTC 443  
Db |||||





QY 121 CTCTACATGGTAACCTAACGACTGACAGTAAACGCTAGTATCGGTATGAGCGCGGATATT 180  
Db 182 CTCATATGCTTTACCAACGACTGTTCCATATAGCAATATTTGTTACGAGGCGGATGAGCTT 241  
QY 181 ATCCCTCACTTACCTGGCTGTGTCCTGCGCTAGCTGTGCGCAATACATCAAGATGCTGG 240  
Db 242 ATTCTGCACACACCGCGGTGATACCTTTGTGTCAGGACGCGCAATACATCAAGTGTGCG 301  
QY 241 ATCCCTGTGAGCCCVACCGTGCCTGAAGTCGCGCTGCGCGCCGACCGGCTCTCTCCGC 300  
Db 302 ACCCCAGTGACACTACAGTGGGAGTCAAGTACGTCGAGCAACCCGCTTCGATACGC 361  
QY 301 AGCAGCGGATATGATGGTGGGCGCGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360  
Db 362 AGTCATGTGACCTATTAGTGGCGCGCCAGGATGTGCTCTGCGCTCTACGTGGGTGAC 421  
QY 361 CTTTGTGGAGCGCTATTCTTGTGTGGCAGGGGTCTCATGGAGACATCCCGAGCAATTGG 420  
Db 422 ATGTGTGGGCTGTCTTCTCTGTGGGCAAGCCCTTCACGTTTCAGACCTGTGCGCCATCAA 481  
QY 421 ACTGTCCAGGACTGCAACTGTTC 443  
Db 482 ACGTCCAGACTGTAACGTCTC 504

RESULT 7  
US-09-878-281A-25  
; Sequence 25, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Imogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph  
; FILE REFERENCE: 35  
; CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 25  
; LENGTH: 541  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-878-281A-25

Query Match 52.1%; Score 233; DB 4; Length 541;  
Best Local Similarity 70.2%; Pred. No. 2.7e-60;  
Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;  
QY 1 GACCGAATTAATTTGCAACAGGGAATTTACCTGGTGTGCTCTTCTCTATCTTCTCTTCTG 60  
Db 62 GACGGATAACTTCGCAACAGGGAATTTGCGCGGTGCTCTTTCTATCTTCTCTTCTTCT 121  
QY 61 GCTTTGTCTCATGCTTGTCTTACCCACAGCGCGGCTGGAGTACCGTAATGCCCTCCGGA 120  
Db 122 GCTCTGTCTTCTTGTCTTAAATCCATCCAGCAGCTGTCTAGAGTGGGGAATACGCTGGC 181  
QY 121 CTCTACATGGTAACACTGACGTACCGTATGCTGTATGAGCGCGGGATATT 180  
Db 182 CTCATGTCTTACCAACGACTGTTCCATATAGTAGTATTGTATGAGCGCGATACGCTT 241  
QY 181 ATCCCTCCACTTACCTGGCTGTGCTTACCCACAGCGCGGCTGGAGTACCGTAATGCCCTCCG 240  
Db 242 ATCTGCACGCGCGGCTGTGCTGTACCTTGTGTCAGGACGGCAATACGCTACATGCTGG 301  
QY 241 ATCCCTGTGAGCCCVACCGTGCCTGGAAGTGGCGCGCCCTGCGCGCCACCGCTCTCTCCGC 300  
Db 302 ACCCCAGTAACTTACAGTGGCAGTACGTCGCGGCAACCCAGCTTCGATACGC 361  
QY 301 AGCAGCTGGATATGATGGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
Db 362 AGTCATGTGACCTGTTAGTGGCGCGCCAGATGTGCTCTGCGCTTACGTGGGTGAT 421  
QY 361 CTTTGTGGAGCGCTATTCTTGTGGGCGAGGGTCTCATGGAGACATCGCCAGCAATTGG 420

Db 422 ATGTGTGGGCGCTTCTCTGTGGGCAACGCTTACGTTTCAGACCCCGCGCATCAA 481  
QY 421 ACTGTCCAGGACTGCAACTGTTC 443  
Db 482 ACGTCCAGACTGTAACGTCTC 504  
RESULT 8  
US-09-878-281A-27  
; Sequence 27, Application US/09878281A  
; Patent No. 6762024  
; GENERAL INFORMATION:  
; APPLICANT: Imogenetics N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph  
; FILE REFERENCE: 35  
; CURRENT APPLICATION NUMBER: US/09/878,281A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 27  
; LENGTH: 541  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-878-281A-27

Query Match 52.1%; Score 233; DB 4; Length 541;  
Best Local Similarity 70.2%; Pred. No. 2.7e-60;  
Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;  
QY 1 GACCGAATTAATTTGCAACAGGGAATTTACCTGGTGTGCTTCTCTATCTTCTCTTCTG 60  
Db 62 GACGGATAACTTCGCAACAGGGAATTTGCGCGGTGCTCTTTCTATCTTCTCTTCTTCT 121  
QY 61 GCTTTGTCTCATGCTTGTCTTACCCACAGCGCGGCTGGAGTACCGTAATGCCCTCCGGA 120  
Db 122 GCTCTGTCTTCTTGTCTTAAATCCATCCAGCAGCTGTCTAGAGTGGGGAATACGCTGGC 181  
QY 121 CTCTACATGGTAACACTGACGTACCGTATGCTGTATGAGCGCGGGATATT 180  
Db 182 CTCATGTCTTACCAACGACTGTTCCATATAGTAGTATTGTATGAGCGCGATGACGTT 241  
QY 181 ATCCCTCCACTTACCTGGCTGTGCTTACCCACAGCTGCTGCGCAATACATCAAGATCTGG 240  
Db 242 ATCTGCACGCGCGGCTGTGCTGTACCTTGTGTCAGGACGGCAATACGCTACATGCTGG 301  
QY 241 ATCCCTGTGAGCCCVACCGTGCCTGGAAGTGGCGCGCCCTGCGCGCCACCGCTCTCTCCGC 300  
Db 302 ACCCCAGTAACTTACAGTGGCAGTACGTCGCGGCAACCCAGCTTCGATACGC 361  
QY 301 AGCAGCTGGATATGATGGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
Db 362 AGTCATGTGACCTGTTAGTGGCGCGCCAGATGTGCTCTGCGCTTACGTGGGTGAT 421  
QY 361 CTTTGTGGAGCGCTATTCTTGTGGGCGAGGGTCTCATGGAGACATCGCCAGCAATTGG 420  
Db 422 ATGTGTGGGCGCTCTTCTCTGTGGGCAAGCCCTTCACGTTTCAGACCCCGCGCATCAA 481  
QY 421 ACTGTCCAGGACTGCAACTGTTC 443  
Db 482 ACGTCCAGACTGTAACGTCTC 504

RESULT 9  
US-08-612-973-29  
; Sequence 29, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE

;; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
;; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
;; NUMBER OF SEQUENCES: 111  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NIXON & VANDERHYE P.C.  
;; STREET: 1100 NORTH GLEBE ROAD  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: U.S.A.  
;; ZIP: 22201-4714  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/612,973  
;; FILING DATE: 11-MAR-1996  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BYRNE, THOMAS E.  
;; REGISTRATION NUMBER: 32,205  
;; REFERENCE/DOCKET NUMBER: 1487-10  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 816-4000  
;; TELEFAX: (703) 816-4100  
;; INFORMATION FOR SEQ ID NO: 29:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 630 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..627  
;; FEATURE:  
;; NAME/KEY: mat\_peptide  
;; LOCATION: 1..624  
US-08-612-973-29

Query Match 52.1%; Score 233; DB 3; Length 630;  
Best Local Similarity 70.2%; Pred. No. 2.9e-60;  
Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;  
QY 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTTGCTCTTCTCTATCTTCTCTG 60  
DB 124 GACGGGATAAATTCGCAACAGGGAATTTGCCCGGTGCTCTTCTATTTCTCTCTC 183  
QY 61 GCTTTGTTCTCATGCTTGTACACCAAGCGGGCTGGAGTACCGTAATCGCTCCGGA 120  
DB 184 GCTCTGTTCTCTTCTTAAATTCATCCAGCAGTAGCTAGAGTGGCGGAATACGCTGCG 243  
QY 121 CTCATCATGTAACCTAACAGCTCAGTACCGTAGTATCGTGTATGAGCGCGGGATATT 180  
DB 244 CTCATGTCCTTACCAACAGCTGTTCCAAATAGCAGTATTGTGTACGAGCGCGATGACGTT 303  
QY 181 ATCTCTCCACTTACCTGGCTGTGTCCTCGTACGCTCTGGCAATACATCAAGATGCTGG 240  
DB 304 ATCTGCAACACCCCGCTGCAATACCTGTGTCCAGGACGGCAATACATCCAGTGTGG 363  
QY 241 ATCCCTGTGAGCCVACCGTCCGGTGAAGTCCCGTGGCCGCGCAACCGCTCTCTCCGC 300  
DB 364 ACCCCAGTGACACCTACAGTGGCAGTCAAGTAGCTGGAGCAACCAACCGCTTCGATACG 423  
QY 301 AGCAGCTGGATATGATGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGAGAC 360  
DB 424 AGTCATGTGGACCTATTAGTGGGCGCGCCAGCATGTGCTCTCGCTCTACGTGGGTGAC 483  
QY 361 CTTTGTGAGCGCTATTTCTTCTGTGGCAGGGGTTCTCATGTGAGACATCGCCAGATTGG 420

DB 484 ATGTGTGGGCTGTCTTCTCTGGGACAAGCCTTCAGCTTCAGACCTCGTCCCATCAA 543  
QY 421 ACTGTCCAGAGCTGCAACTGTTTC 443  
DB 544 ACGGTCCAGACCTGTAACCTGCTC 566  
RESULT 10  
US-08-927-597-29  
;; Sequence 29, Application US/08927597  
;; Patent No. 6245503  
;; GENERAL INFORMATION:  
;; APPLICANT: MAERTENS, GEERT  
;; APPLICANT: BOSMAN, FONS  
;; APPLICANT: DE MARTYNOFF, GUY  
;; APPLICANT: BUYSE, MARIE-ANGE  
;; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
;; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
;; NUMBER OF SEQUENCES: 111  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NIXON & VANDERHYE P.C.  
;; STREET: 1100 NORTH GLEBE ROAD  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: U.S.A.  
;; ZIP: 22201-4714  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/927,597  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/612,973  
;; FILING DATE: 11-MAR-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BYRNE, THOMAS E.  
;; REGISTRATION NUMBER: 32,205  
;; REFERENCE/DOCKET NUMBER: 1487-10  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 816-4000  
;; TELEFAX: (703) 816-4100  
;; INFORMATION FOR SEQ ID NO: 29:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 630 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..627  
;; FEATURE:  
;; NAME/KEY: mat\_peptide  
;; LOCATION: 1..624  
US-08-927-597-29

Query Match 52.1%; Score 233; DB 3; Length 630;  
Best Local Similarity 70.2%; Pred. No. 2.9e-60;  
Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;  
QY 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTTGCTCTTCTCTATCTTCTCTG 60  
DB 124 GACGGGATAAATTCGCAACAGGGAATTTGCCCGGTGCTCTTCTATTTCTCTCTC 183  
QY 61 GCTTTGTTCTCATGCTTGTACACCAAGCGGGCTGGAGTACCGTAATCGCTCCGGA 120  
DB 184 GCTCTGTTCTCTTCTTAAATTCATCCAGCAGTAGCTAGAGTGGCGGAATACGCTGCG 243



```
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 574
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-09-878-281A-120

Query Match      48.3%; Score 215.8; DB 4; Length 574;
Best Local Similarity 67.3%; Pred. No. 4.2e-55;
Matches 301; Conservative 2; Mismatches 144; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTCTA 159
Db 100 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTCTA 120
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCAAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db 160 GCACCTTCTCTCGTGTGACTGTCCGGTGTGCTCTTCTATCTCTCTCTCTCTA 219
QY 121 CTCTACATGTAACGACGTAACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTAT 180
Db 220 ATTTATCAGTCAACCAATGACTGCGGAATCTTAGTATAGTATAGTATAGTATAGTAT 279
QY 181 ATCTCTCACTTACCTGGCTGTGCTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240
Db 280 ATCATGCACTTACCGAGGTGTGCTGCTTGGTGTGAGAACCGGGAACACCTCGCTCTGG 339
QY 241 ATCCCTGTGAGCCVACCGTCCCGTGAAGTCCCGTCCCGTCCCGTCCCGTCTCTCCGC 300
Db 340 GTTCTTTAAACCCCACTGTGCTGCCCTTATGTGGCGCGCGCTCGAATCCATCGCG 399
QY 301 ACGCAGTGGATATGATGTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 400 CGGACGTGGACTTAAATGTGGTGGTGGCGCACCGTCTGCTCGGCCCTGTACATCGGAGAC 459
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGAGGGGTTCCTCATGGAGACATCGCCAGCATTTGG 420
Db 460 CTTTGGGAGGTGTCTTCTGCTGGGAGAGATGTTTACCTTCGGCGCGCGCCATTGG 519
QY 421 ACTGTCAGGAGTGCACATGTTCCATC 447
Db 520 ACTACCCAGGACTGCAACTGCTCTATC 546

RESULT 15
US-09-878-281A-181
; Sequence 181, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 181
; LENGTH: 579
; TYPE: DNA
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152)..(152)
; OTHER INFORMATION: "n" is any nucleotide
US-09-878-281A-181

Query Match      48.3%; Score 215.8; DB 4; Length 579;
Best Local Similarity 67.5%; Pred. No. 4.2e-55;
Matches 301; Conservative 1; Mismatches 144; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 60
Db 100 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 159
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCAAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db 160 GCACCTTCTCTCGTGTGACTGTCCGGTGTGCTCTTCTATCTCTCTCTCTCTA 219
QY 121 CTCTACATGTAACGACGTAACGAGTACGAGTACGAGTACGAGTACGAGTAT 180
Db 220 ATTTATCAGTCAACCAATGACTGCGGAATCTTAGTATAGTATAGTATAGTATAGTAT 279
QY 181 ATCTCTCACTTACCTGGCTGTGCTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240
Db 280 ATCATGCACTTACCGAGGTGTGCTGCTTGGTGTGAGAACCGGGAACACCTCGCTCTGG 339
QY 241 ATCCCTGTGAGCCVACCGTCCCGTGAAGTCCCGTCCCGTCCCGTCTCTCCGC 300
Db 340 GTTCTTTAAACCCCACTGTGCTGCCCTTATGTGGGCGCGCGCTCGAATCCATCGCG 399
QY 301 ACGCAGTGGATATGATGTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 400 CGGACGTGGACTTAAATGTGGTGGTGGCGCACCGTCTGCTCGGCCCTGTACATCGGAGAC 459
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGAGGGGTTCCTCATGGAGACATCGCCAGCATTTGG 420
Db 460 CTTTGGGAGGTGTCTTCTGCTGGGAGAGATGTTTACCTTCGGCGCGCGCCATTGG 519
QY 421 ACTGTCAGGAGTGCACATGTTCCATC 447
Db 520 ACTACCCAGGACTGCAACTGCTCTATC 546

RESULT 14
US-09-878-281A-175
; Sequence 175, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 579
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-09-878-281A-175

Query Match      48.3%; Score 215.8; DB 4; Length 579;
Best Local Similarity 67.3%; Pred. No. 4.2e-55;
Matches 301; Conservative 2; Mismatches 144; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 60
Db 100 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 159
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCAAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db 160 GCACCTTCTCTCGTGTGACTGTCCGGTGTGCTCTTCTATCTCTCTCTCTCTA 219
QY 121 CTCTACATGTAACGACGTAACGAGTACGAGTACGAGTACGAGTACGAGTAT 180
Db 220 ATTTATCAGTCAACCAATGACTGCGGAATCTTAGTATAGTATAGTATAGTATAGTAT 279
QY 181 ATCTCTCACTTACCTGGCTGTGCTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240
Db 280 ATCATGCACTTACCGAGGTGTGCTGCTTGGTGTGAGAACCGGGAACACCTCGCTCTGG 339
QY 241 ATCCCTGTGAGCCVACCGTCCCGTGAAGTCCCGTCCCGTCCCGTCTCTCCGC 300
Db 340 GTTCTTTAAACCCCACTGTGCTGCCCTTATGTGGGCGCGCGCTCGAATCCATCGCG 399
QY 301 ACGCAGTGGATATGATGTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 400 CGGACGTGGACTTAAATGTGGTGGTGGCGCACCGTCTGCTCGGCCCTGTACATCGGAGAC 459
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGAGGGGTTCCTCATGGAGACATCGCCAGCATTTGG 420
Db 460 CTTTGGGAGGTGTCTTCTGCTGGGAGAGATGTTTACCTTCGGCGCGCGCCATTGG 519
QY 421 ACTGTCAGGAGTGCACATGTTCCATC 447
Db 520 ACTACCCAGGACTGCAACTGCTCTATC 546

RESULT 13
US-09-878-281A-175
; Sequence 175, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 579
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-09-878-281A-175

Query Match      48.3%; Score 215.8; DB 4; Length 579;
Best Local Similarity 67.3%; Pred. No. 4.2e-55;
Matches 301; Conservative 2; Mismatches 144; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 60
Db 100 GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTGTGCTCTTCTATCTCTCTCTG 159
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCAAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db 160 GCACCTTCTCTCGTGTGACTGTCCGGTGTGCTCTTCTATCTCTCTCTCTCTA 219
QY 121 CTCTACATGTAACGACGTAACGAGTACGAGTACGAGTACGAGTACGAGTAT 180
Db 220 ATTTATCAGTCAACCAATGACTGCGGAATCTTAGTATAGTATAGTATAGTATAGTAT 279
QY 181 ATCTCTCACTTACCTGGCTGTGCTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240
Db 280 ATCATGCACTTACCGAGGTGTGCTGCTTGGTGTGAGAACCGGGAACACCTCGCTCTGG 339
QY 241 ATCCCTGTGAGCCVACCGTCCCGTGAAGTCCCGTCCCGTCCCGTCTCTCCGC 300
Db 340 GTTCTTTAAACCCCACTGTGCTGCCCTTATGTGGGCGCGCGCTCGAATCCATCGCG 399
QY 301 ACGCAGTGGATATGATGTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 400 CGGACGTGGACTTAAATGTGGTGGTGGCGCACCGTCTGCTCGGCCCTGTACATCGGAGAC 459
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGAGGGGTTCCTCATGGAGACATCGCCAGCATTTGG 420
Db 460 CTTTGGGAGGTGTCTTCTGCTGGGAGAGATGTTTACCTTCGGCGCGCGCCATTGG 519
QY 421 ACTGTCAGGAGTGCACATGTTCCATC 447
Db 520 ACTACCCAGGACTGCAACTGCTCTATC 546
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2005, 11:33:49 ; Search time 19001 Seconds  
(without alignments)  
960.365 Million cell updates/sec

Title: US-09-851-138C-51

Perfect score: 447

Sequence: 1 gacgaataaattcgcaac.....aggactgcaactgttccatc 447

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main.\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/PCTUS3 COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US06 COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US07 COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US09 COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pna/US10 COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pna/US11 COMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pna/US12 COMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pna/US13 COMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pna/US14 COMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pna/US15 COMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pna/US16 COMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pna/US17 COMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pna/US18 COMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pna/US19 COMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pna/US20 COMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pna/US21 COMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pna/US22 COMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pna/US23 COMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pna/US24 COMB.seq.\*
- 23: /cgn2\_6/ptodata/1/pna/US25 COMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pna/US26 COMB.seq.\*
- 25: /cgn2\_6/ptodata/1/pna/US27 COMB.seq.\*
- 26: /cgn2\_6/ptodata/1/pna/US28 COMB.seq.\*
- 27: /cgn2\_6/ptodata/1/pna/US29 COMB.seq.\*
- 28: /cgn2\_6/ptodata/1/pna/US30 COMB.seq.\*
- 29: /cgn2\_6/ptodata/1/pna/US31 COMB.seq.\*
- 30: /cgn2\_6/ptodata/1/pna/US32 COMB.seq.\*
- 31: /cgn2\_6/ptodata/1/pna/US33 COMB.seq.\*
- 32: /cgn2\_6/ptodata/1/pna/US34 COMB.seq.\*
- 33: /cgn2\_6/ptodata/1/pna/US35 COMB.seq.\*
- 34: /cgn2\_6/ptodata/1/pna/US36 COMB.seq.\*
- 35: /cgn2\_6/ptodata/1/pna/US37 COMB.seq.\*
- 36: /cgn2\_6/ptodata/1/pna/US38 COMB.seq.\*
- 37: /cgn2\_6/ptodata/1/pna/US39 COMB.seq.\*
- 38: /cgn2\_6/ptodata/1/pna/US40 COMB.seq.\*
- 39: /cgn2\_6/ptodata/1/pna/US41 COMB.seq.\*
- 40: /cgn2\_6/ptodata/1/pna/US42 COMB.seq.\*
- 41: /cgn2\_6/ptodata/1/pna/US43 COMB.seq.\*
- 42: /cgn2\_6/ptodata/1/pna/US44 COMB.seq.\*
- 43: /cgn2\_6/ptodata/1/pna/US45 COMB.seq.\*

- 44: /cgn2\_6/ptodata/1/pna/US099F COMB.seq.\*
- 45: /cgn2\_6/ptodata/1/pna/US099G COMB.seq.\*
- 46: /cgn2\_6/ptodata/1/pna/US100A COMB.seq.\*
- 47: /cgn2\_6/ptodata/1/pna/US100B COMB.seq.\*
- 48: /cgn2\_6/ptodata/1/pna/US101A COMB.seq.\*
- 49: /cgn2\_6/ptodata/1/pna/US101B COMB.seq.\*
- 50: /cgn2\_6/ptodata/1/pna/US102A COMB.seq.\*
- 51: /cgn2\_6/ptodata/1/pna/US102B COMB.seq.\*
- 52: /cgn2\_6/ptodata/1/pna/US103A COMB.seq.\*
- 53: /cgn2\_6/ptodata/1/pna/US103B COMB.seq.\*
- 54: /cgn2\_6/ptodata/1/pna/US104A COMB.seq.\*
- 55: /cgn2\_6/ptodata/1/pna/US104B COMB.seq.\*
- 56: /cgn2\_6/ptodata/1/pna/US105A COMB.seq.\*
- 57: /cgn2\_6/ptodata/1/pna/US105B COMB.seq.\*
- 58: /cgn2\_6/ptodata/1/pna/US106A COMB.seq.\*
- 59: /cgn2\_6/ptodata/1/pna/US107A COMB.seq.\*
- 60: /cgn2\_6/ptodata/1/pna/US107B COMB.seq.\*
- 61: /cgn2\_6/ptodata/1/pna/US107C COMB.seq.\*
- 62: /cgn2\_6/ptodata/1/pna/US107D COMB.seq.\*
- 63: /cgn2\_6/ptodata/1/pna/US108A COMB.seq.\*
- 64: /cgn2\_6/ptodata/1/pna/US108B COMB.seq.\*
- 65: /cgn2\_6/ptodata/1/pna/US109A COMB.seq.\*
- 66: /cgn2\_6/ptodata/1/pna/US109B COMB.seq.\*
- 67: /cgn2\_6/ptodata/1/pna/US109C COMB.seq.\*
- 68: /cgn2\_6/ptodata/1/pna/US110 COMB.seq.\*
- 69: /cgn2\_6/ptodata/1/pna/US6000 COMB.seq.\*
- 70: /cgn2\_6/ptodata/1/pna/US6001 COMB.seq.\*
- 71: /cgn2\_6/ptodata/1/pna/US6002 COMB.seq.\*
- 72: /cgn2\_6/ptodata/1/pna/US6003 COMB.seq.\*
- 73: /cgn2\_6/ptodata/1/pna/US6004 COMB.seq.\*
- 74: /cgn2\_6/ptodata/1/pna/US6005 COMB.seq.\*
- 75: /cgn2\_6/ptodata/1/pna/US6006 COMB.seq.\*
- 76: /cgn2\_6/ptodata/1/pna/US6007 COMB.seq.\*
- 77: /cgn2\_6/ptodata/1/pna/US6008 COMB.seq.\*
- 78: /cgn2\_6/ptodata/1/pna/US6009 COMB.seq.\*
- 79: /cgn2\_6/ptodata/1/pna/US6010 COMB.seq.\*
- 80: /cgn2\_6/ptodata/1/pna/US6011 COMB.seq.\*
- 81: /cgn2\_6/ptodata/1/pna/US6012 COMB.seq.\*
- 82: /cgn2\_6/ptodata/1/pna/US6013 COMB.seq.\*
- 83: /cgn2\_6/ptodata/1/pna/US6014 COMB.seq.\*
- 84: /cgn2\_6/ptodata/1/pna/US6015 COMB.seq.\*
- 85: /cgn2\_6/ptodata/1/pna/US6016 COMB.seq.\*
- 86: /cgn2\_6/ptodata/1/pna/US6017 COMB.seq.\*
- 87: /cgn2\_6/ptodata/1/pna/US6018 COMB.seq.\*
- 88: /cgn2\_6/ptodata/1/pna/US6019 COMB.seq.\*
- 89: /cgn2\_6/ptodata/1/pna/US6020 COMB.seq.\*
- 90: /cgn2\_6/ptodata/1/pna/US6021 COMB.seq.\*
- 91: /cgn2\_6/ptodata/1/pna/US6022 COMB.seq.\*
- 92: /cgn2\_6/ptodata/1/pna/US6023A COMB.seq.\*
- 93: /cgn2\_6/ptodata/1/pna/US6023B COMB.seq.\*
- 94: /cgn2\_6/ptodata/1/pna/US6024 COMB.seq.\*
- 95: /cgn2\_6/ptodata/1/pna/US6025 COMB.seq.\*
- 96: /cgn2\_6/ptodata/1/pna/US6026 COMB.seq.\*
- 97: /cgn2\_6/ptodata/1/pna/US6027 COMB.seq.\*
- 98: /cgn2\_6/ptodata/1/pna/US6028 COMB.seq.\*
- 99: /cgn2\_6/ptodata/1/pna/US6029 COMB.seq.\*
- 100: /cgn2\_6/ptodata/1/pna/US6030 COMB.seq.\*
- 101: /cgn2\_6/ptodata/1/pna/US6031 COMB.seq.\*
- 102: /cgn2\_6/ptodata/1/pna/US6032 COMB.seq.\*
- 103: /cgn2\_6/ptodata/1/pna/US6033 COMB.seq.\*
- 104: /cgn2\_6/ptodata/1/pna/US6034 COMB.seq.\*
- 105: /cgn2\_6/ptodata/1/pna/US6035 COMB.seq.\*
- 106: /cgn2\_6/ptodata/1/pna/US6036 COMB.seq.\*
- 107: /cgn2\_6/ptodata/1/pna/US6037 COMB.seq.\*
- 108: /cgn2\_6/ptodata/1/pna/US6038 COMB.seq.\*
- 109: /cgn2\_6/ptodata/1/pna/US6039 COMB.seq.\*
- 110: /cgn2\_6/ptodata/1/pna/US6040 COMB.seq.\*
- 111: /cgn2\_6/ptodata/1/pna/US6041 COMB.seq.\*
- 112: /cgn2\_6/ptodata/1/pna/US6042 COMB.seq.\*
- 113: /cgn2\_6/ptodata/1/pna/US6043 COMB.seq.\*
- 114: /cgn2\_6/ptodata/1/pna/US6044 COMB.seq.\*
- 115: /cgn2\_6/ptodata/1/pna/US6045 COMB.seq.\*
- 116: /cgn2\_6/ptodata/1/pna/US6046 COMB.seq.\*

117: /cgn2\_6/ptodata/1/pna/US6047\_COMB.seq.\*  
 118: /cgn2\_6/ptodata/1/pna/US6048\_COMB.seq.\*  
 119: /cgn2\_6/ptodata/1/pna/US6049\_COMB.seq.\*  
 120: /cgn2\_6/ptodata/1/pna/US6050\_COMB.seq.\*  
 121: /cgn2\_6/ptodata/1/pna/US6051\_COMB.seq.\*  
 122: /cgn2\_6/ptodata/1/pna/US6052\_COMB.seq.\*  
 123: /cgn2\_6/ptodata/1/pna/US6053\_COMB.seq.\*  
 124: /cgn2\_6/ptodata/1/pna/US6054\_COMB.seq.\*  
 125: /cgn2\_6/ptodata/1/pna/US6055\_COMB.seq.\*  
 126: /cgn2\_6/ptodata/1/pna/US6056\_COMB.seq.\*  
 127: /cgn2\_6/ptodata/1/pna/US6057\_COMB.seq.\*  
 128: /cgn2\_6/ptodata/1/pna/US6058\_COMB.seq.\*  
 129: /cgn2\_6/ptodata/1/pna/US6059\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	445.8	99.7	447	14	US-08-836-075A-51
2	445.8	99.7	447	14	US-08-836-075B-51
3	445.8	99.7	447	21	US-09-374-494A-51
4	445.8	99.7	447	21	US-09-374-494A-51
5	445.8	99.7	447	36	US-09-851-138B-51
6	445.8	99.7	447	36	US-09-851-138B-51
7	445.8	99.7	447	36	US-09-851-138C-51
8	445.8	99.7	447	36	US-09-851-138C-51
9	255.8	57.2	957	14	US-08-836-075B-220
10	255.8	57.2	957	21	US-09-374-494A-220
11	255.8	57.2	957	36	US-09-851-138B-220
12	255.8	57.2	957	36	US-09-851-138C-220
13	252.2	56.4	957	14	US-08-836-075B-218
14	252.2	56.4	957	21	US-09-374-494A-218
15	252.2	56.4	957	36	US-09-851-138B-218
16	252.2	56.4	957	36	US-09-851-138C-218
17	238.2	53.3	957	14	US-08-836-075B-219
18	238.2	53.3	957	21	US-09-374-494A-219
19	238.2	53.3	957	36	US-09-851-138B-219
20	236.6	52.9	957	14	US-08-836-075B-219
21	236.6	52.9	957	21	US-09-374-494A-217
22	236.6	52.9	957	36	US-09-851-138B-217
23	236.6	52.9	957	36	US-09-851-138C-217
24	236.2	52.8	541	9	US-08-362-455-13
25	236.2	52.8	541	9	US-08-362-455-17
26	236.2	52.8	541	28	US-09-638-693-13
27	236.2	52.8	541	28	US-09-638-693-17
28	236.2	52.8	541	37	US-09-873-224-13
29	236.2	52.8	541	37	US-09-873-224-17
30	236.2	52.8	541	37	US-09-873-224-19
31	236.2	52.8	541	37	US-09-873-224A-13
32	236.2	52.8	541	37	US-09-873-224A-17
33	236.2	52.8	541	38	US-09-899-046-13
34	236.2	52.8	541	38	US-09-899-046-17
35	236.2	52.8	541	38	US-09-899-046A-13
36	236.2	52.8	541	38	US-09-899-046A-17
37	234.6	52.5	541	9	US-08-362-455-19
38	234.6	52.5	541	28	US-09-638-693-19
39	234.6	52.5	541	37	US-09-873-224-19
40	234.6	52.5	541	37	US-09-873-224A-19
41	234.6	52.5	541	38	US-09-899-046-19
42	234.6	52.5	541	38	US-09-899-046A-19
43	234.6	52.5	957	14	US-08-836-075B-215
44	234.6	52.5	957	21	US-09-374-494A-215
45	234.6	52.5	957	36	US-09-851-138B-215
			957	36	US-09-851-138C-215

## ALIGNMENTS

RESULT 1  
 US-08-836-075A-51  
 ; Sequence 51, Application US/08836075A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MAERTENS, GERT  
 ; STUYVER, LIEVEN  
 ; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
 ; AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
 ; AGENTS  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSER: ARNOLD, WHITE & DURKEE  
 ; STREET: P.O. BOX 4433  
 ; CITY: HOUSTON  
 ; STATE: TEXAS  
 ; COUNTRY: USA  
 ; ZIP: 77210-4433  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Microsoft Word 6.0 / ASCII text output  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/836,075A  
 ; FILING DATE: 21-Apr-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/EP95/04155  
 ; FILING DATE: 23 Oct 1995  
 ; APPLICATION NUMBER: EP 94870166.9  
 ; FILING DATE: 21 Oct 1994  
 ; APPLICATION NUMBER: EP 95870076.7  
 ; FILING DATE: 28 Jun 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KAMMERER, PATRICIA A.  
 ; REGISTRATION NUMBER: 29,775  
 ; REFERENCE/DOCKET NUMBER: INNS:004  
 ; INFORMATION FOR SEQ ID NO: 51:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 447 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 51:  
 ; US-08-836-075A-51

Query Match 99.7%; Score 445.8; DB 14; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACGGAATTAAATTCGCAACAGGGAATTTACCTGGTGTCTTCTTCTATCTTCTTCTG 60  
 Db 1 GACGGAATTAAATTCGCAACAGGGAATTTACCTGGTGTCTTCTTCTATCTTCTTCTG 60  
 QY 61 GCTTTGTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
 Db 61 GCTTTGTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
 QY 121 CTCTACATGGTAACCACTGACGTAAGTATGATGATGATGATGATGATGATGATGATGAT 180  
 Db 121 CTCTACATGGTAACCACTGACGTAAGTATGATGATGATGATGATGATGATGATGATGAT 180  
 QY 181 ATCTCTCACTTACCTGGCTGTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Db 181 ATCTCTCACTTACCTGGCTGTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 241 ATCCCTGTGAGCCCTACCGTCCGCTGAGTCCCTGCGCCGCGCCACCGCTCTCTCCGC 300  
 Db 241 ATCCCTGTGAGCCCTACCGTCCGCTGAGTCCCTGCGCCGCGCCACCGCTCTCTCCGC 300  
 QY 301 ACGCAGTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360



Db 301 ACCACGTCGATATGATGGTGGGCGGCACCTATGCTCAGCTCTACGTAGGAGAC 360  
QY 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
Db 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
Db 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

## RESULT 2

US-08-836-075B-51  
; Sequence 51, Application US/08836075B  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; FILE OF INVENTION: prophylactic, therapeutic and diagnostic agents.  
; FILE REFERENCE: 2551-105  
; CURRENT APPLICATION NUMBER: US/08/836, 075B  
; CURRENT FILING DATE: 1997-04-21  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 51  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-08-836-075B-51

Query Match 99.7%; Score 445.8; DB 14; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
Db 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
Db 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
QY 121 CTCTACATGTAATTAACGACTGCACTAACGGTAGTATCGTGTATGAGCGCGGGATATT 180  
Db 121 CTCTACATGTAATTAACGACTGCACTAACGGTAGTATCGTGTATGAGCGCGGGATATT 180  
QY 181 ATCTCTCACTTACCTGGTGTGTCCTCGGTACGCTCTGGCAATACATCAAGATGCTGG 240  
Db 181 ATCTCTCACTTACCTGGTGTGTCCTCGGTACGCTCTGGCAATACATCAAGATGCTGG 240  
QY 241 ATCCCTGTGAGCCYACCGTGGCGGTGAAGTCCGCTGCGCCGACCGCTCTCTCCGC 300  
Db 241 ATCCCTGTGAGCCYACCGTGGCGGTGAAGTCCGCTGCGCCGACCGCTCTCTCCGC 300  
QY 301 AGCAGCTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
Db 301 AGCAGCTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
QY 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
Db 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
Db 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

## RESULT 3

US-09-374-494-51

; Sequence 51, Application US/09374494  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: STUYVER, LIEVEN  
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS  
; TITLE OF INVENTION: GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD WHITE & DURKEE  
; CITY: HOUSTON  
; STATE: TX  
; COUNTRY: US  
; ZIP: 77057  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/374,494  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAMMERER, PATRICIA A.  
; REGISTRATION NUMBER: 29,775  
; REFERENCE/DOCKET NUMBER: INNU:004--1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713.787.1400  
; TELEFAX: 713.787.1440  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-374-494-51

Query Match 99.7%; Score 445.8; DB 21; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
Db 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
Db 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
QY 121 CTCTACATGTAATTAACGACTGCACTAACGGTAGTATCGTGTATGAGCGCGGGATATT 180  
Db 121 CTCTACATGTAATTAACGACTGCACTAACGGTAGTATCGTGTATGAGCGCGGGATATT 180  
QY 181 ATCTCTCACTTACCTGGTGTGTCCTCGGTACGCTCTGGCAATACATCAAGATGCTGG 240  
Db 181 ATCTCTCACTTACCTGGTGTGTCCTCGGTACGCTCTGGCAATACATCAAGATGCTGG 240  
QY 241 ATCCCTGTGAGCCYACCGTGGCGGTGAAGTCCGCTGCGCCGACCGCTCTCTCCGC 300  
Db 241 ATCCCTGTGAGCCYACCGTGGCGGTGAAGTCCGCTGCGCCGACCGCTCTCTCCGC 300  
QY 301 AGCAGCTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
Db 301 AGCAGCTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
QY 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
Db 361 CTTTGTGGAGCGCTATTCTTGTGGCAGGGGTTCTCATGGAGACATCCAGCATTTGG 420  
QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

Db 421 ACTGTCAGGACTGCAACTGTTCCATC 447

RESULT 4  
US-09-374-494A-51  
; Sequence 51, Application US/09374494A  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; FILE REFERENCE: 46 USd1v1  
; CURRENT APPLICATION NUMBER: US/09/374,494A  
; CURRENT FILING DATE: 1999-08-13  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
; US-09-374-494A-51

Query Match 99.7%; Score 445.8; DB 21; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GACGGAATTAATTTCGCAACAGGGAATTTACCTGGTTGCTTTCTCTATCTTCTCTCTG	60
Db	1	GACGGAATTAATTTCGCAACAGGGAATTTACCTGGTTGCTTTCTCTATCTTCTCTCTG	60
QY	61	GCCTTTCTCTCAGCTTGCTTACACCCACAGCGCGGCTGGAGTACCGTAATGCTCCGGA	120
Db	61	GCCTTTGTTCTCAGCTTGCTTACACCCACAGCGCGGCTGGAGTACCGTAATGCTCCGGA	120
QY	121	CTCTACATGGTAAGTAACGACTGACGTAAACGGTAGTATCGTGTATGAGCCGGGATATT	180
Db	121	CTCTACATGGTAAGTAACGACTGACGTAAACGGTAGTATCGTGTATGAGCCGGGATATT	180
QY	181	ATCCTCCACTTACCTGGCTGTGCCCTGGTACGCTCTGGCAATACATCAAGATGCTGG	240
Db	181	ATCCTCCACTTACCTGGCTGTGCCCTGGTACGCTCTGGCAATACATCAAGATGCTGG	240
QY	241	ATCCCTGTGAGCCGACCCGTGCGCGTGAAGTGCCTCGCGCCGACCGCTCTCTCCGC	300
Db	241	ATCCCTGTGAGCCGACCCGTGCGCGTGAAGTGCCTCGCGCCGACCGCTCTCTCCGC	300
QY	301	AGCAGCTGGATATGATGTTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGAGAC	360
Db	301	AGCAGCTGGATATGATGTTGGGCGGCGCACCCCTATGCTCAGCTCTCTACGTAGAGAC	360
QY	361	CTTTGTGGAGCGCTATTTCTCTGCGGAGGGGTTCTCATGAGACATCGCCAGCATGG	420
Db	361	CTTTGTGGAGCGCTATTTCTCTGCGGAGGGGTTCTCATGAGACATCGCCAGCATGG	420
QY	421	ACTGTCAGGACTGCAACTGTTCCATC 447	
Db	421	ACTGTCAGGACTGCAACTGTTCCATC 447	

RESULT 5  
US-09-851-138-51  
; Sequence 51, Application US/09851138  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; STUDYER, LIEVEN  
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
; AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:

/	ADDRESSEE:	ARNOLD, WHITE & DURKEE
/	STREET:	P.O. BOX 4433
/	CITY:	HOUSTON
/	STATE:	TEXAS
/	COUNTRY:	USA
/	ZIP:	77210-4433
/	COMPUTER READABLE FORM:	
/	MEDIUM TYPE:	Floppy disk
/	COMPUTER:	IBM PC compatible
/	OPERATING SYSTEM:	PC-DOS/MS-DOS
/	SOFTWARE:	Microsoft Word 6.0 / ASCII text output
/	CURRENT APPLICATION DATA:	
/	APPLICATION NUMBER:	US/09/851,138
/	FILING DATE:	09-May-2001
/	PRIOR APPLICATION DATA:	
/	APPLICATION NUMBER:	08/836,075
/	FILING DATE:	<Unknown>
/	APPLICATION NUMBER:	EP 94870166.9
/	FILING DATE:	21 Oct 1994
/	APPLICATION NUMBER:	EP 95870076.7
/	FILING DATE:	28 Jun 1995
/	ATTORNEY/AGENT INFORMATION:	
/	NAME:	KAMMERER, PATRICIA A.
/	REGISTRATION NUMBER:	29, 75
/	REFERENCE/DOCKET NUMBER:	INNS:004
/	INFORMATION FOR SEQ ID NO: 51:	
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	447 base pairs
/	TYPE:	nucleic acid
/	STRANDEDNESS:	single
/	TOPOLOGY:	linear
/	MOLECULE TYPE:	cdna
/	HYPOTHETICAL:	NO
/	ANTI-SENSE:	NO
/	SEQUENCE DESCRIPTION:	SEQ ID NO: 51:
/	US-09-851-138-51	
Qy	Query Match	99.7%; Score 445.8; DB 36; Length
Dd	Best Local Similarity	100.0%; Pred. No. 1.5e-114;
	Matches 447; Conservative	0; Mismatches 0; Indels
Qy	1	GACGGAAATTAATTTCGCAACAGGGAAATTTACCTGGTGTCTCTTTCTCTCTCCTA
Dd	1	GACGGAAATTAATTTCGCAACAGGGAAATTTACTCGTGTTGCCTTTCTCTCTCTC
Qy	61	GCTTTGTTCTCATGCTTGCTTACACCACAGCCGGGCTGGAGTAGACCGTATA
Dd	61	GCTTTGTTCTCATGCTTGCTTACACCACAGCCGGCTGGAGTAGACCGTATA
Qy	121	CTCTACATGGTAACGACTGCAGTAAACGTAAGTATCGTGTATGAGAGG
Dd	121	CTCTACATGGTAACGAACGACTGCAGTAAACGTAAGTATCGTGTATGAGAGG
Qy	181	ATCCTCCACTTACCTGGCTGTGCCCTGCGTACGCTCTGGCAATACATAT
Dd	181	ATCCTCCACTTACCTGGCTGTGCCCTGCGTACGCTCTGGCAATACATAT
Qy	241	ATCCCTGTGAGCCCACCGTGCCTCGCTGAAGTCGCCCTGCGCGCACCGG
Dd	241	ATCCCTGTGAGCCCACCGTGCCTCGCTGAAGTCGCCCTGCGCGCACCGG
Qy	301	ACGCAGTGGAATATGATGGTGGRGGGCCACCCTATGCTAGCTCTCTCT
Dd	301	ACGCAGTGGAATATGATGGTGGRGGGCCACCCTATGCTAGCTCTCTCT
Qy	361	CTTTGTGGAGCGCTATTTCTTGTVGGGAGGGGTTCTCATGGAGACATC
Dd	361	CTTTGTGGAGCGCTATTTCTTGTVGGGAGGGGTTCTCATGGAGACATC
Qy	421	ACTGTCCAGGACTGCAACTGTFTCCATC 447
Dd	421	ACTGTCCAGGACTGCAACTGTFTCCATC 447

RESULT 6  
US-09-851-138B-51  
; Sequence 51, Application US/09851138B  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; FILE REFERENCE: 2551-106  
; CURRENT APPLICATION NUMBER: US/09/851,138B  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 51  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-851-138B-51

Query Match 99.7%; Score 445.8; DB 36; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTCCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
DB 1 GACGGAATTAATTTCCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60

QY 61 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
DB 61 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120

QY 121 CTCTACATGTAATAACGACTGACGTAACGAGTATGCTGTATGAGCCGGGGATATT 180  
DB 121 CTCTACATGTAATAACGACTGACGTAACGAGTATGCTGTATGAGCCGGGGATATT 180

QY 181 ATCTCTCACTTACCTGGCTGTGTCCTGCGTACGCTCTGCGCAATACATCAAGATCTGG 240  
DB 181 ATCTCTCACTTACCTGGCTGTGTCCTGCGTACGCTCTGCGCAATACATCAAGATCTGG 240

QY 241 ATCCCTGTGAGCCYACCGTCCGCGTGAAGTCCGCTGCGCGCCACCGCTCTCTCCGC 300  
DB 241 ATCCCTGTGAGCCYACCGTCCGCGTGAAGTCCGCTGCGCGCCACCGCTCTCTCCGC 300

QY 301 AGCAGCTGGATATGATGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
DB 301 AGCAGCTGGATATGATGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGCGAGGGGTTCTATGAGACATGCCAGCATTTGG 420  
DB 361 CTTTGTGGAGCGCTATTTCTTGTGGGCGAGGGGTTCTATGAGACATGCCAGCATTTGG 420

QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
DB 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

RESULT 7  
US-09-851-138C-51  
; Sequence 51, Application US/09851138C  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; FILE REFERENCE: 2551-106  
; CURRENT APPLICATION NUMBER: US/09/851,138C  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7

Query Match 99.7%; Score 445.8; DB 36; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTCCGCAACAGGGAATTTACCTGGTGTCTCTTCTCTATCTCTCTCTG 60  
DB 1 GACGGAATTAATTTCCGCAACAGGGAATTTACCTGGTGTCTTCTCTATCTCTCTCTCTG 60

QY 61 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
DB 61 GCTTTGTTCTCATGCTTGTCTTACCCACAGCGGGCTGGAGTACCGTAATGCTCCGGA 120

QY 121 CTCTACATGTAATAACGACTGACGTAACGAGTATGCTGTATGAGCCGGGGATATT 180  
DB 121 CTCTACATGTAATAACGACTGACGTAACGAGTATGCTGTATGAGCCGGGGATATT 180

QY 181 ATCTCTCACTTACCTGGCTGTGTCCTGCGTACGCTCTGCGCAATACATCAAGATCTGG 240  
DB 181 ATCTCTCACTTACCTGGCTGTGTCCTGCGTACGCTCTGCGCAATACATCAAGATCTGG 240

QY 241 ATCCCTGTGAGCCYACCGTCCGCGTGAAGTCCGCTGCGCGCCACCGCTCTCTCCGC 300  
DB 241 ATCCCTGTGAGCCYACCGTCCGCGTGAAGTCCGCTGCGCGCCACCGCTCTCTCCGC 300

QY 301 AGCAGCTGGATATGATGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360  
DB 301 AGCAGCTGGATATGATGTGGGCGCGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGCGAGGGGTTCTATGAGACATGCCAGCATTTGG 420  
DB 361 CTTTGTGGAGCGCTATTTCTTGTGGGCGAGGGGTTCTATGAGACATGCCAGCATTTGG 420

QY 421 ACTGTCCAGGACTGCAACTGTTCCATC 447  
DB 421 ACTGTCCAGGACTGCAACTGTTCCATC 447

RESULT 8  
US-08-836-075B-220  
; Sequence 220, Application US/08836075B  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; FILE REFERENCE: 2551-105  
; CURRENT APPLICATION NUMBER: US/08/836,075B  
; CURRENT FILING DATE: 1997-04-21  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 220  
; LENGTH: 957  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-08-836-075B-220

Query Match 57.2%; Score 255.8; DB 14; Length 957;  
Best Local Similarity 73.6%; Pred. No. 6e-61;  
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTCCGCAACAGGGAATTTACCTGGTGTCTTCTCTATCTCTCTCTG 60

```
Db 478 GACGGAATAAACTTTGCAACAGGGAAATTTGCCGGTTGCTCCCTTTCTAFTCTCTTCTT 537
QY 61 GCTTTGTTCTCATGCTTTGCTTACACCAACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120
Db 538 GCTCTCTCTCATGCTTTGGTTTCCCGCCAGCGGGGCTAGAGTACAGGAACACGTCGGC 597
QY 121 CTCTACATGGTAACAACTGACGTAACGGTAGTATCGTGTATGAGGCGGGGATATT 180
Db 598 CTATACATACTTACCAACGACTGCTTAAACAGAGCATCGTGTATGAGGCTGATAATGTC 657
QY 181 ATCTCCACTTACTGCTGTGTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 658 ATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 241 ATCCCTGTGAGCCYACCGTCCGGTGAAGTCCCTGCGCCGCGCCGCTCTCTCCGC 300
Db 718 ACGCAGTAACACCGACAGTGGTGTCAAAACATCTTGGCGCAGTCAACGATCAATCCGC 777
QY 301 ACGCACGTGATATGATGTTGGGCGGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 778 AGCATGTGATTTGATGTTGGTGCAGCCACCGCTGTGTTACGACTCTATGTTGGAGAT 837
QY 361 CTTTGGAGCGCTATTTCTTGTGAGGAGGCTTCTCATGAGAGACATCCCGAGCATTTGG 420
Db 838 TTGTGGCGGCTGTTTCTTGTGGCCAAAGCGTTCACCTTTCAGAGCTCGGCAACATTAT 897
QY 421 ACTGTCCAGGACTGCNACTGTTTC 443
Db 898 ACCGTCCAGTTGTGCAATTGCTC 920
```

## RESULT 9

```
US-09-374-494A-220
; Sequence 220, Application US/09374494A
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as
; FILE REFERENCE: 46 usdiv1
; CURRENT APPLICATION NUMBER: US/09/374,494A
; CURRENT FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: EP 94870166.9
; PRIOR FILING DATE: 1994-10-21
; PRIOR APPLICATION NUMBER: EP 95870076.7
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 302
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 220
; LENGTH: 957
; TYPE: DNA
; ORGANISM: hepatitis C virus
```

```
US-09-374-494A-220
Query Match 57.2%; Score 255.8; DB 21; Length 957;
Best Local Similarity 73.6%; Pred. No. 6e-61;
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 1 GACGGAATTAATTTGCAACAGGGAATTTACCTGTTGCTCTTTCTCTATCTTCTCTCTG 60
Db 478 GACGGAATAAACTTTGCAACAGGGAATTTGCCGGTTGCTCCCTTTCTAFTCTCTTCTT 537
QY 61 GCTTTGTTCTCATGCTTTGCTTACACCAACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120
Db 538 GCTCTCTCTCATGCTTTGGTTTCCCGCCAGCGGGCTAGAGTACAGGAACACGTCGGC 597
QY 121 CTCTACATGGTAACAACTGACGTAACGGTAGTATCGTGTATGAGGCGGGGATATT 180
Db 598 CTATACATACTTACCAACGACTGCTTAAACAGAGCATCGTGTATGAGGCTGATAATGTC 657
QY 181 ATCTCCACTTACTGCTGTGTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 658 ATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
```

```
Query Match 57.2%; Score 255.8; DB 21; Length 957;
Best Local Similarity 73.6%; Pred. No. 6e-61;
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 1 GACGGAATTAATTTGCAACAGGGAATTTACCTGTTGCTCTTTCTCTATCTTCTCTCTG 60
Db 478 GACGGAATAAACTTTGCAACAGGGAATTTGCCGGTTGCTCCCTTTCTAFTCTCTTCTT 537
QY 61 GCTTTGTTCTCATGCTTTGCTTACACCAACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120
Db 538 GCTCTCTCTCATGCTTTGGTTTCCCGCCAGCGGGCTAGAGTACAGGAACACGTCGGC 597
QY 121 CTCTACATGGTAACAACTGACGTAACGGTAGTATCGTGTATGAGGCGGGGATATT 180
Db 598 CTATACATACTTACCAACGACTGCTTAAACAGAGCATCGTGTATGAGGCTGATAATGTC 657
QY 181 ATCTCCACTTACTGCTGTGTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 658 ATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
```

```
QY 241 ATCCCTGTGAGCCYACCGTCCCGTGAAGTCCCTGCGCCGCGCACCGCTCTCTCCGC 300
Db 718 ACGCAGTAACACCGACAGTGGTGTCAAAACATCTTGGCGCAGTCAACGATCAATCCGC 777
QY 301 ACGCACGTGATATGATGTTGGGCGGCCACCCCTATGCTCAGCTCTCTCATGTTAGGAGAC 360
Db 778 AGCATGTGATTTGATGTTGGTGCAGCCACCGCTGTGTTACGACTCTATGTTGGAGAT 837
QY 361 CTTTGGAGCGCTATTTCTTGTGAGGAGGCTTCTCATGAGAGACATCCCGAGCATTTGG 420
Db 838 TTGTGGCGGCTGTTTCTTGTGGCCAAAGCGTTCACCTTTCAGAGCTCGGCAACATTAT 897
QY 421 ACTGTCCAGGACTGCNACTGTTTC 443
Db 898 ACCGTCCAGTTGTGCAATTGCTC 920
```

## RESULT 10

```
US-09-851-138B-220
; Sequence 220, Application US/09851138B
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as
; FILE REFERENCE: 2551-106
; CURRENT APPLICATION NUMBER: US/09/851,138B
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: EP 94870166.9
; PRIOR FILING DATE: 1994-10-21
; PRIOR APPLICATION NUMBER: EP 95870076.7
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 302
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 220
; LENGTH: 957
; TYPE: DNA
; ORGANISM: hepatitis C virus
```

```
US-09-851-138B-220
Query Match 57.2%; Score 255.8; DB 36; Length 957;
Best Local Similarity 73.6%; Pred. No. 6e-61;
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 1 GACGGAATTAATTTGCAACAGGGAATTTACCTGTTGCTCTTTCTCTATCTTCTCTCTG 60
Db 478 GACGGAATAAACTTTGCAACAGGGAATTTGCCGGTTGCTCCCTTTCTAFTCTCTTCTT 537
QY 61 GCTTTGTTCTCATGCTTTGCTTACACCAACAGCCGGCTGGAGTACCGTAATGCTCCGGA 120
Db 538 GCTCTCTCTCATGCTTTGGTTTCCCGCCAGCGGGCTAGAGTACAGGAACACGTCGGC 597
QY 121 CTCTACATGGTAACAACTGACGTAACGGTAGTATCGTGTATGAGGCGGGGATATT 180
Db 598 CTATACATACTTACCAACGACTGCTTAAACAGAGCATCGTGTATGAGGCTGATAATGTC 657
QY 181 ATCTCCACTTACTGCTGTGTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 658 ATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 241 ATCCCTGTGAGCCYACCGTCCGGTGAAGTCCCTGCGCCGCGCACCGCTCTCTCCGC 300
Db 718 ACGCAGTAACACCGACAGTGGTGTCAAAACATCTTGGCGCAGTCAACGATCAATCCGC 777
QY 301 ACGCACGTGATATGATGTTGGGCGGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db 778 AGCATGTGATTTGATGTTGGTGCAGCCACCGCTGTGTTACGACTCTATGTTGGAGAT 837
QY 361 CTTTGGAGCGCTATTTCTTGTGAGGAGGCTTCTCATGAGAGACATCCCGAGCATTTGG 420
Db 838 TTGTGGCGGCTGTTTCTTGTGGCCAAAGCGTTCACCTTTCAGAGCTCGGCAACATTAT 897
QY 421 ACTGTCCAGGACTGCNACTGTTTC 443
```

Db  
898 ACCGTCCAGTTGTGCAATTGCTC 920

RESULT 11

```

US-09-851-138C-220
; Sequence 220, Application US/09851138C
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as
; TITLE OF INVENTION: prophylactic, therapeutic and diagnostic agents.
; FILE REFERENCE: 2551-106
; CURRENT APPLICATION NUMBER: US/09851.138C
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: EP 94870166.9
; PRIOR FILING DATE: 1994-10-21
; PRIOR APPLICATION NUMBER: EP 95870076.7
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 302
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 220
; LENGTH: 957
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-09-851-138C-220

```

Query Match 57.2%; Score 255.8; DB 36; Length 957;  
Best Local Similarity 73.6%; Pred. No. 6e-61;  
Matches 326; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy	1	GACGGAAATTAATTCGCAACAGGGAAATTTA	CCTGTTGGCTCTTTCTCTATCTTCCTCTCG	60
Db	478	GACGGAAATAAACTTTGCAACAGGGAAATTTG	CCGGTTGCTCCTTTTCTATCTTCCTCTCTT	537
Qy	61	GCCTTGTCTCATGCTTGCTTACACCCACAG	CCGGCTGGAGTACCGTAATGCCCTCCGGA	120
Db	538	GCTCTCTTCTCATGCTTGTTGTTTCCCCG	CAGGGGCTAGAGTACAGAAACAGTCCGCG	597
Qy	121	CTCTACATGTGTAATAACGACTGCAGTAA	CGGTAGTATCGTGTATGAGCCCGGGATATT	180
Db	598	CTATACATACTTACCAACGACTGCTCTAAC	AGCAGCATCGTGTATGAGGCTGATAATGTC	657
Qy	181	ATCTCCACACTTACCTGGCTGTCCCTGCGT	AGCTCTGGCAATACATCAAGATGCTGG	240
Db	658	ATCTTGACATGCCCGGCTGTGGCTTGCACT	TCGCGAGGGTAACCACTCAAGGTGCTGG	717
Qy	241	ATCCCTGTGAGCCCYACCGTCCGCGTGAAGT	CGCCCTGCGCGCCACGCGCTCTCTCCGC	300
Db	718	ACGCCAGTAACACCGACAGTGGCTGTCAAA	CATCCTGGCGCAGTCACCGCATCAATCCGC	777
Qy	301	ACGCACGTGATATGATGGTGGRGCGGCCAC	CCCTATGCTCAGCTCTCTACGTAGGAGAC	360
Db	778	AGGCATGTGAATTTGATGGTGGGTGCAACC	ACGCTGTGTTGAGCACTCTATGTTGGAGAT	837
Qy	361	CTTTGTGGAGCGCTATTCTTTGTGGGCAGG	GGTTCTCATGGAGACATCGCCAGCATGG	420
Db	838	TTGTGCGGGCTGTTTTCTTGTGGGCCA	GGTTTCATTTTCAGAGCTCGGCAACATTAT	897
Qy	421	ACTGTCCAGACTGCAACTGTTC	443	
Db	898	ACGCTCCAGTTGTCAATTGCTC	920	

## RESULT 12

```

RESOLI 12
US-08-836-075B-218
/ Sequence 218, Application US/08836075B
/ GENERAL INFORMATION:
/ APPLICANT: INNOGENETICS N.V.
/ TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as
/ FILE OF INVENTION: prophylactic, therapeutic and diagnostic agents.
/ FILE REFERENCE: 2551-105
/ CURRENT APPLICATION NUMBER: US/08/836.075B
/ CURRENT FILING DATE: 1997-04-21

```

```

; PRIORITY NUMBER: EP 94870166.9
; PRIOR APPLICATION NUMBER: EP 94870166.9
; PRIOR FILING DATE: 1994-10-21
; PRIOR APPLICATION NUMBER: EP 95870076.7
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 302
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 957
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-08-836-075B-218

```

Query Match	56.4%	Score 252.2	DB 14	Length 957
Best Local Similarity	72.9%	Pred. No. 6.2e-60		
Matches 323	Conservative	1	Mismatches 119	Indels 0
Qy	1	GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTTGCTTTCTCTATCTATCTTCCTCTG	60	
Db	478	GACGGAAATTAATTCGCAACAGGGAATTTACCTGGTTGCTTTCTCTATCTATCTTCCTCTG	537	
Qy	61	GCCTTTGTTCTCATGCTTGTCTTACACCCACAGCGGGCTGGAGTACCGTAATGCCTCCGGA	120	
Db	538	GCCTTTGCTCTCTGCTTGGTCTGCTCTGCTGCAAGGATTTGAATACCGGAATGTGCTCGGC	597	
Qy	121	CTCTACATGGTAACCTAACGACTCAGCTAACCGGTAGTATCGTGTATGAGCGCGGGATATT	180	
Db	598	CTCTACGTGCTCACCAACGACTGCTCTAAACCGCAGTATCGTGTATGAGGCCCTCGAATGC	657	
Qy	181	ATCCTCAGCTTACTGCTGCTGTCCCTCGGTACGCTCGGCAATACATCAAGATGCTGG	240	
Db	658	ATCTTGACATTTGCACGGTTGTGTGCCCTGGCTTCAATCAGGCAACTCTCTGCAATGCTGG	717	
Qy	241	ATCCCTGTAGCCCYACCGTCCGGTGAAGTGCCTCGCCGCCACCGGCTCTCTCCGC	300	
Db	718	ATTCCGGTGGCACCAACAGTGGGGGTTAAGTACGCTGGCGCAGCACTGCATCGATCCGC	777	
Qy	301	ACGCAGTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC	360	
Db	778	AGTCATGTGGATCTGCTGGTGGGAGCTGCTACGTTTGTCTCCGCGCTGTATGTTGGCGAT	837	
Qy	361	CTTTGTGGAGCGCTATTTCTTGTGTGGCAGGGGTTCTCATGGAGACATCGCCAGCATTTG	420	
Db	838	ATGTGTGGAGCCGCTCTCTTGTGTGGACAGGCTTTCACCTTCAGACCTCGTCAGCACAC	897	
Qy	421	ACTGTCCAGACTGCAACTGTTTC	443	
Db	898	ACGGTGCAGACCTGCAATTGCTC	920	

## RESULT 13

```

RESULTS: 13
US-09-374-494A-218
; Sequence 218, Application US/09374494A
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as
; ; TITLE OF INVENTION: prophylactic, therapeutic and diagnostic agents.
; FILE REFERENCE: 46 USdiv1
; CURRENT APPLICATION NUMBER: US/09/374,494A
; CURRENT FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: EP 94870166.9
; PRIOR FILING DATE: 1994-10-21
; PRIOR APPLICATION NUMBER: EP 95870076.7
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 302
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 957
; TYPE: DNA
; ORGANISM: hepatitis C virus
US-09-374-494A-218

Query Match 56.4%; Score 252.2; DB 21; Length 957;
Best Local Similarity 72.9%; Pred. No. 6.2e-60;

```

Matches 323; Conservative 1; Mismatches 119; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTTACCTGGTTGCTCTTCTCTATCTCTCTCTCTCTG 60  
Db 478 GACGGAATTAATTTTCGCAACAGGGAATTTTACCGGTTGCTCTTCTCTATCTCTCTCTCTG 537  
QY 61 GCTTTGCTG 120  
Db 538 GCTTTGCTG 597  
QY 121 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 180  
Db 598 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 657  
QY 181 ATCCCTCCACTTACCTGGCTGTGCTCCCTGGTGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 240  
Db 658 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 717  
QY 241 ATCCCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 300  
Db 718 ATCCGCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 777  
QY 301 ACGCAGTGGATATGATGGTGGGCGGCGCCACCGCTCTCTCTAGCTCTCTCTAGTGGAGAC 360  
Db 778 AGTCATGTAATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 837  
QY 361 CTCTGAGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 420  
Db 838 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 897  
QY 421 ACTGTCAGGACTGCAACTGTTTC 443  
Db 898 ACGGTGCAGACCTGCAATTGCTC 920

RESULT 14  
US-09-851-138B-218  
; Sequence 218, Application US/09851138B  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; TITLE OF INVENTION: prophylactic, therapeutic and diagnostic agents.  
; FILE REFERENCE: 2551-106  
; CURRENT APPLICATION NUMBER: US/09/851,138B  
; PRIOR FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 218  
; LENGTH: 957  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-851-138B-218

Query Match 56.4%; Score 252.2; DB 36; Length 957;  
Best Local Similarity 72.9%; Pred. No. 6.2e-60;  
Matches 323; Conservative 1; Mismatches 119; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTTACCTGGTTGCTCTTCTCTATCTCTCTCTCTG 60  
Db 478 GACGGAATTAATTTTCGCAACAGGGAATTTTACCGGTTGCTCTTCTCTATCTCTCTCTCTG 537  
QY 61 GCTTTGCTG 120  
Db 538 GCTTTGCTG 597  
QY 121 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 180  
Db 598 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 657  
QY 181 ATCCCTCCACTTACCTGGCTGTGCTCCCTGGTGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 240  
Db 658 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 717  
QY 241 ATCCCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 300  
Db 718 ATCCGCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 777  
QY 301 ACGCAGTGGATATGATGGTGGGCGGCGCCACCGCTCTCTCTAGCTCTCTCTAGTGGAGAC 360  
Db 778 AGTCATGTAATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 837  
QY 361 CTCTGAGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 420  
Db 838 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 897  
QY 421 ACTGTCAGGACTGCAACTGTTTC 443  
Db 898 ACGGTGCAGACCTGCAATTGCTC 920

QY 181 ATCTCCACTTACCTGGCTGTGCTCCCTGGTGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 240  
Db 658 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 717  
QY 241 ATCCCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 300  
Db 718 ATCCGCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 777  
QY 301 ACGCAGTGGATATGATGGTGGGCGGCGCCACCGCTCTCTCTAGCTCTCTCTAGTGGAGAC 360  
Db 778 AGTCATGTAATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 837  
QY 361 CTCTGAGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 420  
Db 838 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 897  
QY 421 ACTGTCAGGACTGCAACTGTTTC 443  
Db 898 ACGGTGCAGACCTGCAATTGCTC 920

RESULT 15  
US-09-851-138C-218  
; Sequence 218, Application US/09851138C  
; GENERAL INFORMATION:  
; APPLICANT: INNOGENETICS N.V.  
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes and their use as  
; TITLE OF INVENTION: prophylactic, therapeutic and diagnostic agents.  
; FILE REFERENCE: 2551-106  
; CURRENT APPLICATION NUMBER: US/09/851,138C  
; PRIOR FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: EP 94870166.9  
; PRIOR FILING DATE: 1994-10-21  
; PRIOR APPLICATION NUMBER: EP 95870076.7  
; PRIOR FILING DATE: 1995-06-28  
; NUMBER OF SEQ ID NOS: 302  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 218  
; LENGTH: 957  
; TYPE: DNA  
; ORGANISM: hepatitis C virus  
US-09-851-138C-218

Query Match 56.4%; Score 252.2; DB 36; Length 957;  
Best Local Similarity 72.9%; Pred. No. 6.2e-60;  
Matches 323; Conservative 1; Mismatches 119; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTTACCTGGTTGCTCTTCTCTATCTCTCTCTCTG 60  
Db 478 GACGGAATTAATTTTCGCAACAGGGAATTTTACCGGTTGCTCTTCTCTATCTCTCTCTCTG 537  
QY 61 GCTTTGCTG 120  
Db 538 GCTTTGCTG 597  
QY 121 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 180  
Db 598 CTCTACATGTAATTAACGACTGCAAGTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 657  
QY 181 ATCCCTCCACTTACCTGGCTGTGCTCCCTGGTGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 240  
Db 658 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 717  
QY 241 ATCCCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 300  
Db 718 ATCCGCTGGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 777  
QY 301 ACGCAGTGGATATGATGGTGGGCGGCGCCACCGCTCTCTCTAGCTCTCTCTAGTGGAGAC 360  
Db 778 AGTCATGTAATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 837  
QY 361 CTCTGAGAGCCYACCGTGCCTGGAAGTGCCTCGCCGCGCCACCGCTCTCTCTCGG 420  
Db 838 ATCTTGACATCTGCTGGTGGAGTCTGCTTAAACGAGTAAACGAGTAAACGAGTAAACGAGTAAAT 897  
QY 421 ACTGTCAGGACTGCAACTGTTTC 443  
Db 898 ACGGTGCAGACCTGCAATTGCTC 920

Db 838 ATGTGTGGAGCCGTCTTCTTGTGGGACAGGCTTTCACCTTCAGACCTCGTCAGCACAC 897

Qy 421 ACTGTCCAGGACTGCAACTGTTTC 443

Db 898 ACGGTGCAGACCTGCAATTGCTC 920

Search completed: March 4, 2005, 19:52:39  
Job time : 19004 secs



**THIS PAGE BLANK (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2005, 09:20:58 ; Search time 1057 Seconds  
(without alignments)  
2503.429 Million cell updates/sec

Title: US-09-851-138C-51  
Perfect score: 447  
Sequence: 1 gacgaataatttcgaac.....aggactgcaactgtccatc 447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	445.8	99.7	447	2 AAT27962	Aat27962 Hepatitis
2	236.2	52.8	540	2 AAT278031	Aag78031 Hepatitis
3	236.2	52.8	541	2 AAT278029	Aag78029 Hepatitis
4	234.6	52.5	541	2 AAT278033	Aag78033 Hepatitis
5	234.6	52.5	541	2 AAT278036	Aag78036 Hepatitis
6	234.6	52.5	541	2 AAT278032	Aag78032 Hepatitis
7	233	52.1	541	2 AAT278030	Aag78030 Hepatitis
8	233	52.1	541	2 AAT278034	Aag78034 Hepatitis
9	233	52.1	541	2 AAT278035	Aag78035 Hepatitis
10	233	52.1	630	2 AAT12965	Aat12965 HCV E1 co
11	233	52.1	630	6 AAL48929	Aal48929 Hepatitis
12	233	52.1	630	10 ADD5537	Add5537 Hepatitis
13	233	52.1	630	12 ADP71119	Adp71119 HCV DNA e
14	221.4	49.5	1249	2 AAT27858	Aat27858 HCV J1 NS
15	219.4	49.1	565	2 AAT2785072	Aag785072 HCV envel
16	217.4	48.6	579	2 AAT278104	Aag78104 Hepatitis
17	216.2	48.4	580	2 AAT278652	Aaz078652 HCV E dom
18	215.8	48.3	574	2 AAT278081	Aag78081 Hepatitis
19	215.8	48.3	579	2 AAT278106	Aag78106 Hepatitis
20	215.8	48.3	579	2 AAT278103	Aag78103 Hepatitis

21	209.8	46.9	447	2 AAT27956	Aat27956 Hepatitis
22	209	46.8	447	2 AAT27952	Aat27952 Hepatitis
23	208.6	46.7	447	2 AAT27955	Aat27955 Hepatitis
24	206.6	46.2	579	2 AAT278102	Aag78102 Hepatitis
25	206.2	46.1	579	2 AAT278109	Aag78109 Hepatitis
26	205	45.9	579	2 AAT278101	Aag78101 Hepatitis
27	204.6	45.8	574	2 AAT278080	Aag78080 Hepatitis
28	204.2	45.7	1270	2 AAV60668	Aav60668 Fragment
29	204.2	45.7	1562	2 AAV60672	Aav60672 Fragment
30	204.2	45.7	1953	8 AAL55222	Aal55222 Plasmid P
31	204.2	45.7	2116	2 AAT12242	Aat12242 Encodes P
32	204.2	45.7	2829	2 AAV60673	Aav60673 Fragment
33	203	45.4	579	2 AAT278114	Aag78114 Hepatitis
34	203	45.4	579	2 AAT278108	Aag78108 Hepatitis
35	201.4	45.1	579	2 AAT278113	Aag78113 Hepatitis
36	201.4	45.1	579	2 AAT278093	Aag78093 Hepatitis
37	201	45.0	447	2 AAT27951	Aat27951 Hepatitis
38	199.4	44.6	673	2 AAV42305	Aav42305 HCV core
39	199.4	44.6	742	2 AAQ20926	Aaq20926 C10-E15 D
40	199.4	44.6	932	2 AAQ20923	Aaq20923 C10-E12 D
41	198.2	44.3	447	2 AAT27954	Aat27954 Hepatitis
42	197.8	44.3	447	2 AAT27953	Aat27953 Hepatitis
43	196.2	43.9	577	2 AAQ35082	Aaq35082 HCV envel
44	196.2	43.9	633	2 AAT12706	Aat12706 HCV E1 co
45	196.2	43.9	633	6 AAL48915	Aal48915 Hepatitis

## ALIGNMENTS

RESULT 1  
AAT27962  
ID AAT27962 standard; DNA; 447 BP.  
XX  
AC AAT27962;  
DT 11-MAR-1997 (first entry)  
XX  
DE Hepatitis C virus type 10a isolate NN98 bases 478-925.  
XX  
KW Hepatitis C virus; subtype; polymerase chain reaction; amplification;  
KW PCR; primer; probe; antibody; infection; ss.  
XX  
OS Hepatitis C virus.  
XX  
PN WO9613590-A2.  
XX  
PD 09-MAY-1996.  
XX  
PF 23-OCT-1995; 95WO-EP004155.  
XX  
PR 21-OCT-1994; 94EP-00870166.  
PR 28-JUN-1995; 95EP-00870076.  
XX  
(INNO-) INNOGENETICS NV.  
XX  
PI Maertens G, Stuyver L;  
DR WPI; 1996-251460/25.  
DR P-PSDB; AAR96551.  
XX  
PT Hepatitis C virus polynucleic acid unique to unidentified sub-type -  
PT used to develop probes and primers for new subtypes and vaccines to  
PT prevent and treat infection.  
XX  
PS Claim 6; Fig 3; 150pp; English.

XX  
CC The sequences AAT27937-727989 represent novel sequences isolated from  
CC Hepatitis C virus subtypes different from subtypes 1a-c, 2a-d, 3a-f, 4a-  
CC j, 5a and 6a. They esp. from the novel subtypes 1d-f, 2e-1, 2k, 2l, 3g,  
CC 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5'  
CC untranslated region (UR), the Core/E1, NS4 or NS5B regions of the genome.  
CC This sequence represents nucleotides 478-925 from the HCV type 10a

CC isolate NE98. The new HCV types were isolated from patients with chronic  
 CC HCV from the Benelux countries, France, Cameroon and Vietnam, because of  
 CC their aberrant reactivities. The RNA was extracted, cDNA synthesised and  
 CC PCR amplified, cloned and genotyped. The 5'UR, Core/E1 and NS5A regions  
 CC were sequenced either directly or partially and used to classify the new  
 CC viruses into (sub)types based on comparison with known sequences. The  
 CC sequences were used to generate the peptides AAR96424-R96524. The  
 CC sequences can also be used to synthesise probes and primers for the  
 CC detection of HCV in a sample. The polypeptides can be used to detect anti  
 CC -HCV antibodies, for HCV typing or to prevent HCV infections  
 XX

Sequence 447 BP; 82 A; 130 C; 114 G; 118 T; 0 U; 3 Other;

Query Match 99.7%; Score 445.8; DB 2; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-129;  
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTATCTCTCTCTG 60  
 DB 1 GACGGAATTAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTATCTCTCTCTG 60  
 QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
 DB 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
 QY 121 CTCTACATGTAATAACGACTGAGTAACGGTAGTATCGTGTATGAGCCGGGGATATT 180  
 DB 121 CTCTACATGTAATAACGACTGAGTAACGGTAGTATCGTGTATGAGCCGGGGATATT 180  
 QY 181 ATCCTCCACTTACTGGCTGTGTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240  
 DB 181 ATCCTCCACTTACTGGCTGTGTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240  
 QY 241 ATCCCTGTGAGCCYACCGTGCCTGGAAGTGCCTCGCGCCGACCCCTCTCTCCGC 300  
 DB 241 ATCCCTGTGAGCCYACCGTGCCTGGAAGTGCCTCGCGCCGACCCCTCTCTCCGC 300  
 QY 301 ACCGAGTGATATGATGTTGGGGCGGCGCCACCTATGCTCAGCTCTACGTAGGAGAC 360  
 DB 301 ACCGAGTGATATGATGTTGGGGCGGCGCCACCTATGCTCAGCTCTACGTAGGAGAC 360  
 QY 361 CTTTGTGAGCGCTATTCTTGTGGGAGGGGTTCTCATGAGACATCGCCAGCATTTGG 420  
 DB 361 CTTTGTGAGCGCTATTCTTGTGGGAGGGGTTCTCATGAGACATCGCCAGCATTTGG 420  
 QY 421 ACTGTCAGACTGCACTGTTCCATC 447  
 DB 421 ACTGTCAGACTGCACTGTTCCATC 447

## RESULT 2

AAQ78031  
 ID AAQ78031 standard; cDNA; 540 BP.

XX AC AAQ78031;

XX AC AAQ78031;

DT 25-MAR-2003 (revised)

DT 21-JUL-1995 (first entry)

XX Hepatitis C virus Core/E1 region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;

XX classification; immunisation; prophylaxis; serotyping; ss.

OS Hepatitis C virus type 3a.

XX Key Location/Qualifiers

XX CDS 2..541

XX FT /\*tag= a

XX FT /product= "Core/E1 polypeptide."

XX PN W09425601-A2.

PD 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

DR P-PSDB; AAR63279.

XX New polynucleotide sequences from hepatitis C virus - and related  
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,  
 PT treatment, diagnosis and typing of HCV isolates.

XX Claim 2; Page 107-108; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous  
 CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-957 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning  
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.  
 CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated HD10-2-21.  
 CC (Updated on 25-MAR-2003 to correct FN field.)

XX Sequence 540 BP; 105 A; 153 C; 144 G; 138 T; 0 U; 0 Other;

Query Match 52.8%; Score 236.2; DB 2; Length 540;

Best Local Similarity 70.7%; Pred. No. 1.3e-63;

Matches 313; Conservative 1; Mismatches 129; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTATCTCTCTCTG 60  
 DB 61 GACGGAATTAATTTTCGCAACAGGGAATTTACCTGGTGTCTCTTCTATCTCTCTCTG 120  
 QY 61 GCTTTGTTCTCATGCTTGTCTTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120  
 DB 121 GCTCTGTTCTTGTCTTATCCATCCAGCAGCTAGTCTAGAGTGGCGGAAACAGTCTGGC 180  
 QY 121 CTCTACATGTAATAACGACTGAGTAACGGTAGTATCGTGTATGAGCCGGGGATATT 180  
 DB 181 CTCTACATGTAATAACGACTGAGTAACGGTAGTATCGTGTATGAGCCGGGGATATT 240  
 QY 181 ATCCTCCACTTACTGGCTGTGTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 240  
 DB 241 ATCCTCCACTTACTGGCTGTGTCCCTGCTGACGCTCTGGCAATACATCAAGATCTGG 300  
 QY 241 ATCCCTGTGAGCCYACCGTGCCTGGAAGTGCCTCGCGCCGACCCCTCTCTCCGC 300  
 DB 301 ACCGAGTGATATGATGTTGGGGCGGCGCCACCTATGCTCAGCTCTACGTAGGAGAC 360  
 QY 301 ACCGAGTGATATGATGTTGGGGCGGCGCCACCTATGCTCAGCTCTACGTAGGAGAC 360  
 DB 361 AGGCATGTAGACATATTGTTGGGCGGCGCCACGATGTGCTCTCTCTACGTGGTGTAT 420  
 QY 361 CTTTGTGAGCGCTATTCTTGTGGGAGGGGTTCTCATGAGACATCGCCAGCATTTGG 420  
 DB 421 ATGTTGGGGCGGCTCTTCTCTGTTGGGAGCAAGCTTTCACGTTTCAGACCTCGTCCGCAATCAA 480  
 QY 421 ACTGTCAGACTGCACTGTTCCATC 443  
 DB 481 ACCGTCACAGCTGTACTGCTC 503







CC nucleotides selected from an HCV type 3 genomic sequence, more  
 CC particularly (i) the region spanning positions 417-957 of the Core/E1  
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of  
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-  
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype  
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to  
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or  
 CC as a probe for specific detection/classification of nucleic acid.  
 CC Polypeptides encoded by the nucleotides in such compositions may be used  
 CC for immunisation against HCV, for the detection of antibodies directed  
 CC against HCV and for serotyping. This sequence corresponds to the Core/E1  
 CC region of HCV subtype 3a and is taken from a clone designated HD10-2-14.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX

SQ Sequence 541 BP; 106 A; 154 C; 143 G; 138 T; 0 U; 0 Other;

Query Match 52.1%; Score 233; DB 2; Length 541;

Best Local Similarity 70.2%; Pred. No. 1.3e-62;

Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;

QY 1 GACGGATTAATTCGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTTCCTCTG 60  
 Db 62 GACGGATAAATTTGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTCCTCT 121  
 QY 61 GCTTTGTCTCATGCTTGTACACCCAGCCGGCTGGAGTACCGTAATGCTCCGGA 120  
 Db 122 GCTCTGTCTTCTTAATCCATCCAGCAGCTAGTCTAGAGTGGGGAACACGCTGCG 181  
 QY 121 CTCTACATGTAATAACGACTGCTTCCATAGCAGTATGTGTATGAGCCGATGACGTT 180  
 Db 182 CTCTATGCTCTTACCAAGCACTTCCATAGCAGTATGTGTATGAGCCGATGACGTT 241  
 QY 181 ATCTCCATTAATTCGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTTCCTCTG 240  
 Db 242 ATCTCTGCTCTTCTTAATCCATCCAGCAGCTAGTCTAGAGTGGGGAACACGCTGCG 301  
 QY 241 ATCTCTGCTGAGCCCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 Db 302 ACCCAGTACACCTTACCAAGCACTTCCATAGCAGTATGTGTATGAGCCGATGACGTT 361  
 QY 301 ACSCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 Db 362 AGGCATGTAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421  
 QY 361 CTTTGGGAGCGTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420  
 Db 422 ATGTGTGGGCGCTTCTCTCTGTTGGGCAAGCCTTTCAGCTTCAGACCTCTGTCGCATCAA 481  
 QY 421 ACTGTCAGGACTGCAACTGTTTC 443  
 Db 482 ACGGTCCAGACCTGTAACTGCTC 504

# RESULT 8

AAQ78034

ID AAQ78034 standard; cDNA; 541 BP.

XX AAQ78034;

XX 25-MAR-2003 (revised)

DT 01-AUG-1995 (first entry)

XX Hepatitis C virus Core/E1 region.

DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;

KW classification; immunisation; prophylaxis; serotyping; ss.

XX Hepatitis C virus type 3a.

OS Key Location/Qualifiers

XX 2.541

PH /\*tag= a

FT CDS

FT

/product= "Core/E1 polypeptide."

WO9425601-A2.

10-NOV-1994.

27-APR-1994; 94WO-BP001323.

27-APR-1993; 93EP-00401099.

05-AUG-1993; 93EP-00402019.

(INNO-) INNOGENETICS NV SA.

Maertens G, Stuyver L;

WPI; 1994-358277/44.

P-PSDB; AAR63282.

New polynucleotide sequences from hepatitis C virus - and related

vectors, polypeptide(s) and antibodies, useful for immunisation,

treatment, diagnosis and typing of HCV isolates.

Claim 2; Page 113-114; 404pp; English.

Compositions comprising at least 5, and pref. 8 or more contiguous

nucleotides selected from an HCV type 3 genomic sequence, more

particularly (i) the region spanning positions 417-957 of the Core/E1

region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of

the NS3 region of HCV type 3; (iii) the region spanning positions 4892-

5292 of the NS3/4 region of HCV type 3; (iv) the region spanning

positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype

3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to

amplify nucleic acid from an isolate belonging to a specific genotype, or

as a probe for specific detection/classification of nucleic acid.

Polypeptides encoded by the nucleotides in such compositions may be used

for immunisation against HCV, for the detection of antibodies directed

against HCV and for serotyping. This sequence corresponds to the Core/E1

region of HCV subtype 3a and is taken from a clone designated BR33-1-10.

(Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 541 BP; 100 A; 157 C; 148 G; 136 T; 0 U; 0 Other;

Query Match 52.1%; Score 233; DB 2; Length 541;

Best Local Similarity 70.2%; Pred. No. 1.3e-62;

Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;

QY 1 GACGGATTAATTTGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTTCCTCTG 60

Db 62 GACGGATAAATTTGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTCCTCTCT 121

QY 61 GCTTTGTCTCATGCTTGTACACCCAGCCGGCTGGAGTACCGTAATGCTCCGGA 120

Db 122 GCTCTGTCTTCTTAATCCATCCAGCAGCTAGTCTAGAGTGGGGAACACGCTGCG 181

QY 121 CTCTACATGTAATAACGACTGCTTCCATAGCAGTATGTGTATGAGCCGATGACGTT 180

Db 182 CTCTATGCTCTTACCAAGCACTTCCATAGCAGTATGTGTATGAGCCGATGACGTT 241

QY 181 ATCTCCATTAATTTGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTTCCTCTG 240

Db 242 ATCTCTGCTCTTCTTAATCCATCCAGCAGCTAGTCTAGAGTGGGGAACACGCTGCG 301

QY 241 ATCTCTGCTGAGCCCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

Db 302 ACCCAGTAAACCTACAGTGGCAGTACGTCAGTGGGGAACACACCGCTTCGATACGC 361

QY 301 ACSCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360

Db 362 AGTCATGTGGACCTGTGTAGTGGCGGCCACGATGTGTCTCTGCGCTTTACGTGGGTGAT 421

QY 361 CTTTGTGGAGCGCTATTTCTTGTGGGCGAGGGTTCATGAGAGACATCGCCAGCATGG 420

Db 422 ATGTGTGGGCGCTTCTCTCTGTTGGGCAAGCCTTTCAGCTTCAGACCTCTGTCGCATCAA 481







```
XX 15-JAN-2004 (first entry)
XX Hepatitis C virus E1/E2 protein coding sequence #5.
XX Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
XX liver fibrosis; ds; gene.
XX Hepatitis C virus.
OS WO2003051912-A2.
XX
XX 26-JUN-2003.
XX
XX 18-DEC-2002; 2002WO-BF014480.
XX
XX 18-DEC-2001; 2001US-00020510.
XX
XX 16-OCT-2002; 2002US-0418358P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Depla E, Bosman F;
XX
XX WPI: 2003-541632/51.
XX
XX P-PSDB; ADD55538.
XX
XX New hepatitis C virus (HCV) vaccine composition, useful for reducing
XX liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX
XX Example 2; SEQ ID NO 29; 271pp; English.
XX
XX The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
XX liver disease. The vaccine of the invention comprises an HCV E1 or E2
XX protein as an antigen. The HCV vaccine is useful for reducing liver
XX disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
XX present DNA sequence encodes an HCV E1/E2 protein.
XX
XX Sequence 630 BP; 127 A; 175 C; 168 G; 160 T; 0 U; 0 Other;
SQ
Query Match 52.1%; Score 233; DB 10; Length 630;
Best Local Similarity 70.2%; Pred. No. 1.4e-62;
Matches 311; Conservative 1; Mismatches 131; Indels 0; Gaps 0;

QY 1 GACGGAATTAATTCGCAACAGGGAATTTACCTGGTGTCTTTCTCTATCTCTCTTCG 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 GACGGGATAAATTTGCAACAGGGAATTTGCCGGTGTCTCTTTCTATTTTCTCTCTC 183
QY 61 GCTTTGTTCTCATGCTGTGTACACCCAGCGGGCTGGAGTACCGTAATGCTCCGGA 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 GCTCTGTTCTCTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGGGAATACGTC 243
QY 121 CTCTCATGTAATTAACGACTGCAGTAACGGTAGTATGCTGTATGAGCGCGGGATATT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 CTCTATGTCCTTACCAACAGCTGTTCATAGCAGTATTTGTACGAGGCCATGACGTT 303
QY 181 ATCTCCCATTAACCTGGCTGTGTCCCTCGGTAACGCTCTGGCAATACATCAAGATCTGG 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
304 ATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGGAGGCAATACATCCACGTCGTGG 363
QY 241 ATCCCTGTGAGCCYACCTGCGGTGAAGTGCCTGCGCGCCACCGCCTCTCTCCG 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 ACCCCAGTGACACCTACAGTGGCAGTCAAGTAGCTCGGAGCAACCAACCGCTTCGATACGC 423
QY 301 AGCCAGCTGATATGATGTGGGCGCGCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
424 AGTCATGTGAGCCTATTATGTGGCGCGGCCACGATGTCTCTCGGCTCTACGTGGGTGAC 483
QY 361 CTTTGTGGAGCGCTATTTCTTGTGGCGAGGGGTTCTATGAGAGACATCCGACGATGG 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
484 ATGTGTGGGGCTGTCTTCTCTCGTGGGACAAGCCTTCACGTTCTCAGACCTCGTCCCATCAA 543
QY 421 ACTGTCCAGGACTGCACCTGTTTC ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 544 ACGTCCAGACCTGTAACTGCTC 566
RESULT 13
ADP71119
ID ADP71119 standard; cDNA; 630 BP.
XX
XX AC ADP71119;
XX
XX 23-SEP-2004 (first entry)
XX
XX HCV DNA encoding E1 protein HCC162.
XX
XX Hepatitis C virus; HCV; E1 glycoprotein; E2 glycoprotein; HCV infection;
XX liver disease; liver fibrosis; ss; serum alanine aminotransferase level;
XX steatosis; anti-E2 immunoreactivity; vaccine.
XX
XX Hepatitis C virus; type 3a.
XX
XX US2004126395-A1.
XX
XX 01-JUL-2004.
XX
XX 18-DEC-2002; 2002US-00321798.
XX
XX 18-DEC-2001; 2001US-0453708P.
XX
XX 16-OCT-2002; 2002US-0418358P.
XX
XX (MAER/) MAERTENS G.
XX
XX (DEPL/) DEPLA E.
XX
XX (BOSM/) BOSMAN F.
XX
XX Maertens G, Depla E, Bosman F;
XX
XX WPI: 2004-499089/47.
XX
XX P-PSDB; ADP71120.
XX
XX Use of hepatitis C virus (HCV) vaccine composition for reducing liver
XX disease, serum alanine aminotransferase levels, steatosis, or anti-E2
XX immunoreactivity in the liver of a chronic HCV-infected mammal.
XX
XX Example 2; SEQ ID NO 29; 176pp; English.
XX
XX The invention relates to the use of a hepatitis C virus (HCV) vaccine
XX composition for reducing liver disease (such as liver fibrosis or its
XX progression), serum alanine aminotransferase (ALT) levels, steatosis, or
XX anti-E2 immunoreactivity in the liver of a chronic HCV-infected mammal,
XX or for treating a chronic HCV-infected mammal. The liver disease is
XX reduced by at least 1-2 points according to the overall Ishak score in
XX the HCV-infected mammal. Also included are a method for predicting
XX changes in liver disease in a chronic HCV-infected mammal, a therapeutic
XX HCV vaccine composition (comprising at least one purified or a
XX combination of at least 2 HCV single or specific oligomeric recombinant
XX envelope protein selected from an E1 or E2 protein, a part of E1 and E2
XX proteins, an E1/E2 protein complex formed from purified HCV single or
XX specific oligomeric recombinant E1 or E2 proteins or its parts and
XX optionally a pharmaceutical adjuvant), a composition (comprising at least
XX one E1 or E2 peptide, and optionally, a pharmaceutical adjuvant), an
XX immunogenic HCV composition (or HCV vaccine composition) comprising a
XX recombinant virus allowing expression of at least one HCV recombinant
XX envelope protein (selected from an E1 protein and/or an E2 protein, and
XX their parts, and optionally, a pharmaceutical adjuvant) and an HCV
XX vaccine composition (comprising a recombinant virus allowing expression
XX of at least one HCV recombinant envelope protein chosen from an E1
XX protein and/or an E2 protein, and parts of the E1 and E2 proteins and,
XX optionally, a pharmaceutical adjuvant. The HCV vaccine composition is
XX useful for reducing liver disease (such as liver fibrosis or its
XX progression), serum ALT levels, steatosis, or anti-E2 immunoreactivity in
XX the liver in a chronic HCV-infected mammal, or for treating a chronic HCV
XX -infected mammal, particularly human. The HCV E1 proteins are useful for
XX in vitro monitoring HCV disease or prognosing the response to treatment
XX of patients suffering from HCV infection. The present sequence is an HCV
XX cDNA encoding an E1 protein (or fragment).
```

XX	SQ	Sequence	630 BP; 127 A; 175 C; 168 G; 160 T; 0 U; 0 Other;
		Query Match	52.1%; Score 233; DB 12; Length 630;
		Best Local Similarity	70.2%; Pred. No. 1.4e-62;
		Matches 311; Conservative	1; Mismatches 131; Indels 0; Gaps 0;
QY	1	GAGCGAATTAATTCGCAACGAGGAATTAACCTGGTTGCTCTTCTATCTTCCTCTTCG 60	
Db	124	GACGGGATAAATTCGCAACGAGGAATTCGCCGGTGTCTCTTTTCTATTTCCTCTCTC 183	
QY	61	GCTTTGTTCTCATGCTTCTTACACACACGCGGCTGGATCGTAACTCCCTCCGGA 120	
Db	184	GCTCTGTTCTCTTGTTAAATTCACGACAGCTAGTCTAGAGTGGCGGAATACGTCCTGC 243	
QY	121	CTCTACATGGTAACTAAACGACTGCAGTCAACGGTAGTAGTGTGTATGAGCGCCGGGATATT 180	
Db	244	CTCTATGTCTTACCAACGACTGTTTCCAATAGCAGTATTGTGTACGAGCGCGATGACGTT 303	
QY	181	ATCTCTCACTTACCTGGCTGTGTCCCTCGGTAACGCTCTGGCAATACATCAAGATGCTGG 240	
Db	304	ATTCGCAACACCCCGCTGTGATACCTTGTGTCCAGGACGGCAATACATCCACGTCGTGG 363	
QY	241	ATCCCTGTGAGCCCYACCGTCGCGTGAAGTGCCTCGCGCCGACACCGCTCTCTCCGC 300	
Db	364	ACCCGAGTGACACCTACAGTGGCAGTCAAGTAGCTCGAGCAACACCGCTTCCATAGC 423	
QY	301	ACGCACGTGGATATGATGTGGTGGGCGGCCACCTATGCTCAAGTCTCTACGTAGGAGAC 360	
Db	424	AGTCATGTGGACCTATTAGTGGGCGGGCCACGATGTGCTCTGCGCTCTACGTGGGTGAC 483	
QY	361	CTTTGTGGAGCGCTATTCTTTGTGCGCAGGGGTTCTCATGGAGACATCGCCACGATGG 420	
Db	484	ATGTGTGGGGCTGTCTTCTCTGTGGGCAAGCCCTTACGTTCAGACCTCGTCGCCATCAA 543	
QY	421	ACTGTCCAGGACTGCACACTGTTCC 443	
Db	544	ACGGTCCAGACCTGTGTACTGCTC 566	

## RESULT 14

AAZ07658  
ID AAZ07658 standard; DNA; 1249 BP.

XX  
AC AAZ07658:XX  
DT 20-MAR-2003 (revised)

DT 08-NOV-1999 (first entry)

XX  
DE HCV J1 NS domain consensusHepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH; ...  
XX  
KW

KW HCV infection; vaccine; ds.

Hepatitis C virus.

XX  
PN  
EP939128-A2.

01-SEP-1999 XX PN

[illegible]

17-SEP-1990; 99EP-00101746.

PR 15-SEP-1989; 89US-00408045.  
PR 21-DEC-1989; 89US-00456142.

PR 17-SEP-1990; 90EP-00310149.  
XX

РПА (ОУАА/) ОУА А.  
РПА (СНП) СЧПМ

CHIRON CORP.

PPI Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;  
PPI Cha T, Irvine BD;

WPI: 1999-480843/4

11

XX	New Hepatitis C Virus isolates, useful for diagnosis of hepatitis infections and development of vaccines.
PT	Disclosure; Fig 14; 132pp; English.
XX	
PS	The invention provides two new isolates of hepatitis C virus (HCV), J1 and J7. These two isolates comprise nucleotide and amino acid sequences that are distinct from the HCV isolate HCV-1. The nucleotide sequences may be used to detect non-A, non-B HCV (NANBH) polynucleotides by hybridisation for diagnosis of NANBH infections. They may also be used to screen blood donors, donated blood and blood products for this infection. The isolates may also be used to isolate other naturally occurring variants of the virus. The polypeptides may be used as a vaccine for administration to patients to protect against infection with NANBH. The present sequence represents the NS1 domain consensus sequence of HCV isolates J1 and HCV-1. (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to correct PR field.)
CC	Sequence 1249 BP; 160 A; 282 C; 276 G; 206 T; 0 U; 325 Other;
XX	
QQ	Query Match 49.5%; Score 221.4; DB 2; Length 1249;
XX	Best Local Similarity 54.6%; Pred. No. 8.2e-59;
XX	Matches 244; Conservative 83; Mismatches 120; Indels 0; Gaps 0;
QY	1 GACGGATTAAATTCGCAACAGGGAATTTACCTGGTTGCTCTTCTCTACTCTCTCTCTG 60
DB	68 GACGGCGTGAACATATGCACAGGGAAYTTCYKCYGGTTGCTCTTCTCTACTCTCTCTG 127
QY	61 GCTTTGTCTCATGCTTGCTTGTACACCCACAGCGGGCTGGAGTACCGTAATGCTCCCGGA 120
DB	128 GCTCTGCTCTCYGTGTGACYRSTCCMGCTTCSGCTAYSAAGTGCACACKYSWCSGGG 187
QY	121 CTCTACATGTAACCTAACGACTCAGTAACGCTAGTATCGTGTATGAGGCGGGGATATT 180
DB	188 MTWTACCAVGTCAAAAGYATGCYCAACTCRAGYATTCGTAYGAGCGGCGSGAYGYS 247
QY	181 ATCTCCACTTACTGGCTGTGTCCCTCGGTACGCTTGGGAATACATCAAGATGCTGG 240
DB	248 ATCMTCGAVRCYCCSGGGTGCSTCCYTGCGTTTCGKAGRRCAAYKCTCSMGKTGYTGG 307
QY	241 ATCCCTGTGAGCCYACCGTCGCGGTGAAGTCCCTGCGCGCCGACCGGCTCTCTCCGC 300
DB	308 GTAGCGMTSACYCYAGSTSGCSRCAAGRATGSCARMTSCCCCKKACGMMRYTWCGA 367
QY	301 ACCGACGTGATATGATGTGGRGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
DB	368 CGYCAKRTCGAYTGTCTGTGGGASSGVCYVYTCGTGTCSCYMTSTACGTGGGGAY 427
QY	361 CTTTGTGGAGCGGTATTTCTTGTGTGGCAGGGGTTCTCATGGAGACATGCCAGCATTTG 420
DB	428 CTTTGGGRTCTGTGTCTTGTCTTCTCCCACTGTTTCCACCTTCTCKCYMGSCGSCAYKRG 487
QY	421 ACTGTCAGGACTGCAACTGTTTCCATC 447
DB	488 ACRRPCARGRYTGCAAYTGCTCWAIC 514
XX	
XX	RESULT 15
ID	AAQ35072
XX	AAQ35072 standard; DNA; 565 BP.
XX	AAQ35072;
XX	
XX	20-MAY-1993 (first entry)
XX	HCV envelope region nucleic acid.
XX	Envelope; region; type C; hepatitis; virus; HCV; vaccine; serum;
KW	non-A, non-B; amplify; ss.
XX	
OS	Hepatitis C virus.
XX	

```
PN JP04349885-A.
XX
PD 04-DEC-1992.
XX
PF 29-MAY-1991; 91JP-00152169.
XX
PR 29-MAY-1991; 91JP-00152169.
XX
PA (TEIJ ) TEIJIN LTD.
XX
XX WPI; 1993-022708/03.
XX
XX Envelope region nucleic acid fragment - for type C hepatitis virus (I),
PT for producing vaccine.
XX
XX Claim 1; Page 2; 13pp; Japanese.
XX
CC This sequence encodes a novel envelope region of type C hepatitis virus
CC (HCV). This fragment can be used for the preparation of a vaccine for
CC hepatitis C. This fragment was prepared from the serum of non-A, non-B
CC hepatitis patients and the envelope region DNA was amplified by PCR using
CC the primer sequences given in AAQ35073-76
XX
SQ Sequence 565 BP; 61 A; 92 C; 106 G; 85 T; 0 U; 221 Other;

Query Match 49.1%; Score 219.4; DB 2; Length 565;
Best Local Similarity 45.2%; Pred. No. 2.5e-58;
Matches 202; Conservative 143; Mismatches 102; Indels 0; Gaps 0;

QY 1 GACGGAAATTAATTTCGACACAGGGAATTACCTGGTTGCTTCTTCTATCTTCTCTCTG 60
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 34 GACRGGGTGACTAYGCAACAGGGAAYTDCYGGTTCYCTTCTATCTTCTCTCTCTG 93
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 61 GCTTTCTTCATGCTTGTTTACACCCACAGCGGCTGGAGTACCGTAATCGCTCCGGA 120
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 94 GCYTBCTVCTGYTGTACRYBCHGYVTCRCYAVVARGTGCRCACAGSSWCMGGG 153
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 121 CTCACATGTTACTAAGTACGACGAGTACGGTAGTATCGTGTATGAGCCGGGATATT 180
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 154 NYRTAYCATGTACBARYGAYTGYYCYAACKRAGYATWGTGYAYRARGYRRMGAYRY 213
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 181 ATCCTCCACTTACTCGCTGTGCTCCCTGCTACGCTCTGGCAATACATCAAGATGCTGG 240
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 214 ATCMTGCWYDYCCBGGGTGYRYBCCYTGTYCGSARRRYRRYNNYTCBMGNTGYTGG 273
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 241 ATCCCTGTGAGCCCYACCGTGCCTGGAAGTCCCTGCGCCGCCACCGCTCTCTCCGC 300
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 274 GYRGCGCTBACVCCACGSGTYGCYACGRRAYVBYAVVYCCRCBRYGVNMTWCGB 333
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 301 ACCGAGTGGATATGATGGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGAC 360
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 334 CGBCAVRTCGAYVTGTYTHGGGRSVGCYRCYCTCTGYTCGSCYVTSTAYTGKGRGAY 393
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 361 CTTTGTGGAGCGCTATTCTTGTGGGCGAGGGTCTCTCATGGAGACATGCCAGATTGG 420
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 394 YBTGCGGRTCYGTYYTYTYRYCYCABYGTGTTTACCYCTCTMBMCYMGSDGYAYGDG 453
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 421 ACTGTCACGACTGCAACTGTTCCATC 447
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 454 ARDVRCAARGVTGYAAATGYTCDMTY 480
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
```

Search completed: March 4, 2005, 11:51:17  
Job time : 1060 secs

THIS PAGE BLANK (USTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2005, 11:12:53 ; Search time 2352 Seconds  
(without alignments)  
7234.153 Million cell updates/sec

Title: US-09-851-138C-51  
Perfect score: 447  
Sequence: 1 gacgaataatttcgaac.....aggactgcaactgttcacatc 447

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gss1:\*
- 9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	44.8	10.0	488	1 AV755731	AV755731
2	42.6	9.5	1517	2 BE636546	BE636546 rockefell
3	42.6	9.5	1559	2 BE636668	BE636668 rockefell
4	42.6	9.5	1561	2 BE636534	BE636534 rockefell
5	39.6	8.9	562	4 BM321463	BM321463 rockefell
6	38.6	8.6	270	2 BE636547	BE636547 rockefell
7	38.6	8.6	409	2 BE636640	BE636640 rockefell
8	38.6	8.6	482	4 BM320876	BM320876 rockefell
9	38.6	8.6	571	2 BE636646	BE636646 rockefell
10	38.6	8.6	650	4 BM321414	BM321414 rockefell
11	38.6	8.6	1416	9 CL961612	CL961612 OeIFCC006
12	38.6	8.6	1538	2 BE636716	BE636716 rockefell
13	37.4	8.4	924	5 BQ707541	BQ707541 AGENCOURT
14	37.2	8.3	458	7 H29400	H29400 ym60b07.r1
15	37.2	8.3	555	4 BM490244	BM490244 ppg2n.pk0
16	37.2	8.3	876	2 BF339031	BF339031 602034820
17	37.2	8.3	643	4 BM321147	BM321147 rockefell
18	36.8	8.2	561	2 BE726729	BE726729 894093B12
19	36.8	8.2	581	2 BE452248	BE452248 894064B06
20	36.8	8.2	631	5 BQ763234	BQ763234 EBr002.SQ
21	36.8	8.2	642	5 BU652417	BU652417 1112099H0
22	36.8	8.2	679	7 CF134176	CF134176 WHE4368.G
23	36.8	8.2	688	2 BE519519	BE519519 HV CB001
24	36.8	8.2	799	2 BF621718	BF621718 HVSMea001

25	36.6	8.2	297	6 C73583	C73583 C73583 Rice
26	36.6	8.2	952	9 CR50009Z	AL060615 Drosophll
C 27	36.2	8.1	753	9 CR280679	CR280679 CR280679
C 28	36.2	8.1	854	5 BX928104	AG055561 Pan trogl
C 29	36.2	8.1	1045	9 AG065561	AG065561 Pan trogl
C 30	36	8.1	481	4 BJ273717	BJ273717 BJ273717
C 31	36	8.1	612	5 BU573228	BU573228 PA_Ea000
C 32	36	8.1	618	6 CA754942	CA754942 BR030090
C 33	36	8.1	637	5 BU574967	BU574967 PA_Ea000
C 34	36	8.1	835	9 CC620321	CC620321 OGLAU65TV
C 35	36	8.1	850	7 CK151620	CK151620 FGAS03425
C 36	36	8.1	865	8 CC339658	CC339658 OGI166TV
C 37	36	8.1	886	9 CC620316	CC620316 OGLAU65TH
C 38	36	8.1	917	9 CG223099	CG223099 CG0FL34TV
C 39	35.8	8.0	277	2 AW922110	AW922110 LG1_236.C
C 40	35.8	8.0	690	9 CL979817	CL979817 OeIFCC044
C 41	35.8	8.0	835	9 CG092989	CG092989 PUIFR16TD
C 42	35.6	8.0	335	9 CC757846	CC757846 ZMMBB014
C 43	35.6	8.0	709	7 CK446695	CK446695 pnc9506AB
C 44	35.4	7.9	548	7 CF615385	CF615385 CES013415
C 45	35.4	7.9	668	9 CL628075	CL628075 OR_CBA004

ALIGNMENTS

RESULT 1  
AV755731/c  
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000  
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.  
ACCESSION AV755731  
VERSION AV755731.1 GI:10913579  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 488)  
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,  
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and  
Chen,Z.  
TITLE Homo sapiens cDNA BM clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES

source  
1..488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BMFAK03"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/clone\_lib="BM"  
/note="Vector: pTriplex2; Site 1: sfIIA; Site 2: sfIIB"

ORIGIN

Query Match 10.0%; Score 44.8; DB 1; Length 488;  
Best Local Similarity 63.8%; Pred. No. 0.061;  
Matches 83; Conservative 1; Mismatches 43; Indels 3; Gaps 1;  
Qy 318 GTGTGGGCGGCGCCATGCTAGCTCTCTACGTAGGACCTTTGTGGACGGCTATT 377  
Db 473 GTGTGTGTACACTCGTCTGCTAGCTCTCTACGTGTGGGACCTCTGCGACGGAGTAT 414

```

QY 378 TCTTGYGGCAGGGTTCTCATGGAGACATCCGACGATTGGACTGTCCAGACTGCAA 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 GCTTGCAGTTTCAGTCATCATCTGG---CCTCAGCACCATGAGTTTGTGCATGAATGCAA 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 438 CTGTTCATC 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 CTGCTCATC 347

RESULT 2
BE636546 1517 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION
rockefeller.0.126 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,
mRNA sequence.
ACCESSION BE636546
VERSION BE636546.2 GI:18055259
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoebidae; Mastigamoeba.
REFERENCE
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
1 (bases 1 to 1517)
Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED 11830664
COMMENT On Aug 25, 2000 this sequence version replaced gi:9919657.
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1517 Std Error: 0.00
POLYA=Yes.
FEATURES
Location/Qualifiers
1..1517
/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
ORIGIN
Query Match 9.5%; Score 42.6; DB 2; Length 1517;
Best Local Similarity 50.8%; Pred. No. 0.33;
Matches 99; Conservative 1; Mismatches 95; Indels 0; Gaps 0;
QY 182 TCCTCCACTTACCTGGCTGTGCCCTGGTAGCTGCGCCGCGCCACCGCTCTCTCCGCA 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1245 TTGCGCAAGTACCCGCTCTGGCGCTTCGCGCTCGGTATGCGCCAGACCGTCGCGC 1304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 TCCTGTGAGCCYACCGTCGCGGTGAGTGCCTTGGCGCGCCACCGCTCTCTCCGCA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1305 TCGGTGTTCATCAAGACCGTGTGCGCAAGGTCCCGCGCAAGACCGGTGCCCGCGGACA 1364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 CGCAGTGGATATGATGTGGRGCGGCCACCGCTATGCTCAGCTCTTACCTAGGAGACC 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1365 AGAAGGACGCCAAGAAGTAGGACCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 1424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 TTGTGGAGCGCTAT 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1425 GTGGACTGCCCTAT 1439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
BE636668 1559 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION
rockefeller.0.341 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,
mRNA sequence.
ACCESSION BE636668
VERSION BE636668.2 GI:18055257
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoebidae; Mastigamoeba.
REFERENCE
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
1 (bases 1 to 1559)
Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED 11830664
COMMENT On Aug 25, 2000 this sequence version replaced gi:9919779.
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1559 Std Error: 0.00
POLYA=Yes.
FEATURES
Location/Qualifiers
1..1559
/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
ORIGIN
Query Match 9.5%; Score 42.6; DB 2; Length 1559;
Best Local Similarity 50.8%; Pred. No. 0.33;
Matches 99; Conservative 1; Mismatches 95; Indels 0; Gaps 0;
QY 182 TCCTCCACTTACCTGGCTGTGCCCTGGTAGCTGCGCCGCGCCACCGCTCTCTCCGCA 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1263 TTGCGCAAGTACCCGCTCTGGCGCTTCGCGCTCGGTATGCGCCAGACCGTCGCGC 1322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 TCCTGTGAGCCYACCGTCGCGGTGAGTGCCTTGGCGCGCCACCGCTCTCTCCGCA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1323 TCGGTGTTCATCAAGACCGTGTGCGCAAGGTCCCGCGCAAGACCGGTGCCCGCGGACA 1382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 CGCAGTGGATATGATGTGGRGCGGCCACCGCTATGCTCAGCTCTTACCTAGGAGACC 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1383 AGAAGGACGCCAAGAAGTAGGACCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 1442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 TTGTGGAGCGCTAT 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1443 GTGGACTGCCCTAT 1457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
BE636534 1561 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION
rockefeller.0.102 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,
mRNA sequence.
ACCESSION BE636534
VERSION BE636534.2 GI:18055196
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoebidae; Mastigamoeba.
REFERENCE
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
1 (bases 1 to 1561)
Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.

```



**TITLE**  
The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

**MEDLINE**  
21819461

**PUBLISHED**  
11830664

**COMMENT**  
On Aug 25, 2000 this sequence version replaced gi:9919645.  
Contact: Muller Miklos  
Laboratory of Biochemical Parasitology  
The Rockefeller University  
1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 1561 Std Error: 0.00  
POLYA=Yes.

**FEATURES**  
Location/Qualifiers  
1..1561  
/organism="Mastigamoeba balamuthi"  
/mol\_type="mRNA"  
/strain="ATCC 30984"  
/db\_xref="taxon:108607"  
/clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"

**ORIGIN**  
Query Match 9.5%; Score 42.6; DB 2; Length 1561;  
Best Local Similarity 50.8%; Pred. No. 0.33; Mismatches 95; Indels 0; Gaps 0;  
Matches 99; Conservative 1;

QY 182 TCCTCCACTTACTGGCTGTGTCCTCGGTAGTCCCTCGGTGCAATACATCAAGATGCTGGA 241  
Db 1261 TTGCGCAAGTACCCGCTCGGCGCTTCCGCGTTCGATATGCGCCAGACCGTGGCG 1320

QY 242 TCCCTGTGAGCCCYACCGTCGCGGTGAAGTCCCTCGGCGCCACCGCTCTCTCCGGA 301  
Db 1321 TCGGTGTATCAAGACCGTCGCGCAAGGTCCCGGCAAGACCGTGGCGCGCGGACA 1380

QY 302 CQACGTGGATATGATGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361  
Db 1381 AGAAGAGCCCAAGAGAAGTGAAGCGCTCTCTGCTTCTGCAACCTCTTAAGAGAC 1440

QY 362 TTGTGGAGCGCTAT 376  
Db 1441 GTGAGTGGCCCTAT 1455

**RESULT 5**  
BM321463  
LOCUS  
DEFINITION  
Mastigamoeba balamuthi lambda ZAP II Library  
Mastigamoeba balamuthi CDNA similar to elongation factor 1 alpha, mRNA sequence.

**ACCESSION**  
BM321463  
VERSION  
BM321463.1 GI:18055869

**KEYWORDS**  
EST.

**SOURCE**  
Mastigamoeba balamuthi

**ORGANISM**  
Mastigamoeba balamuthi

**REFERENCE**  
AUTHORS  
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.  
1 (bases 1 to 562)  
Baptiste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and Philippe, H.

**TITLE**  
The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

**MEDLINE**  
21819461

**PUBLISHED**  
11830664

**COMMENT**  
Contact: Muller Miklos  
Laboratory of Biochemical Parasitology  
The Rockefeller University  
1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 562 Std Error: 0.00  
POLYA=No.

**FEATURES**  
Location/Qualifiers

**TITLE**  
The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

**MEDLINE**  
21819461

**PUBLISHED**  
11830664

**COMMENT**  
On Aug 25, 2000 this sequence version replaced gi:9919645.  
Contact: Muller Miklos  
Laboratory of Biochemical Parasitology  
The Rockefeller University  
1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 1561 Std Error: 0.00  
POLYA=Yes.

**FEATURES**  
Location/Qualifiers  
1..1561  
/organism="Mastigamoeba balamuthi"  
/mol\_type="mRNA"  
/strain="ATCC 30984"  
/db\_xref="taxon:108607"  
/clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"

**ORIGIN**  
Query Match 8.9%; Score 39.6; DB 4; Length 562;  
Best Local Similarity 49.2%; Pred. No. 2.1; Mismatches 96; Conservative 0;

QY 182 TCCTCCACTTACTGGCTGTGTCCTCGGTAGTCCCTCGGTGCAATACATCAAGATGCTGGA 241  
Db 292 TTGCGCAAGTACCCGCTCGGCGCTTCCGCGTTCGATATGCGCCAGACCGTGGCG 351

QY 242 TCCCTGTGAGCCCYACCGTCGCGGTGAAGTCCCTCGGCGCCACCGCTCTCTCCGGA 301  
Db 352 TCGGTGTATCAAGACCGTCGCGCAAGGTCCCGGCAAGACCGTGGCGCGCGGACA 411

QY 302 CQACGTGGATATGATGTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361  
Db 412 AGAAGAGCCCAAGAGAAGTGAAGCGCTCTCTGCTTCTGCAACCTCTTAAGAGAC 471

QY 362 TTGTGGAGCGCTAT 376  
Db 472 GTGAGTGGCCCTAT 486

**RESULT 6**  
BM321463  
LOCUS  
DEFINITION  
Mastigamoeba balamuthi lambda ZAP II Library  
Mastigamoeba balamuthi CDNA similar to elongation factor 1 alpha, mRNA sequence.

**ACCESSION**  
BM321463  
VERSION  
BM321463.1 GI:9919658

**KEYWORDS**  
EST.

**SOURCE**  
Mastigamoeba balamuthi

**ORGANISM**  
Mastigamoeba balamuthi

**REFERENCE**  
AUTHORS  
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.  
1 (bases 1 to 270)  
Baptiste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and Philippe, H.

**TITLE**  
The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

**MEDLINE**  
21819461

**PUBLISHED**  
11830664

**COMMENT**  
Contact: Muller Miklos  
Laboratory of Biochemical Parasitology  
The Rockefeller University  
1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 270 Std Error: 0.00  
POLYA=Yes.

**FEATURES**  
Location/Qualifiers  
1..270  
/organism="Mastigamoeba balamuthi"  
/mol\_type="mRNA"  
/strain="ATCC 30984"  
/db\_xref="taxon:108607"  
/clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"

**ORIGIN**  
Query Match 8.6%; Score 38.6; DB 2; Length 270;  
Best Local Similarity 54.8%; Pred. No. 3.6; Mismatches 74; Conservative 1;

QY 242 TCCCTGTGAGCCCYACCGTGGCGGTGAAGTGGCGCGCCACCGCTCTCTCCGGA 301

Db 28 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 87  
 QY 302 CCACGTGTATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361  
 Db 88 AGAAGGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGCAACCTCTCTAAGAGAGC 147  
 QY 362 TTTGTGGAGCGCTAT 376  
 Db 148 GTGGACTGGCCCTAT 162

RESULT 7  
 BE636640  
 LOCUS  
 DEFINITION  
 Mastigamoeba balamuthi 409 bp mRNA linear EST 03-JAN-2002  
 rockefeller.0.284 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 409)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 409 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..409  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 2; Length 409;  
 Best Local Similarity 54.8%; Pred. No. 3.9;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 152 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 211

QY 302 CCACGTGTATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 212 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 271

QY 362 TTTGTGGAGCGCTAT 376

Db 272 GTGGACTGGCCCTAT 286

RESULT 8  
 BM320876  
 LOCUS  
 DEFINITION  
 rockefeller.0.139 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 482 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..482  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 482;  
 Best Local Similarity 54.8%; Pred. No. 4;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 243 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 302

QY 302 CGCACGTGGATATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 303 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 362

QY 362 TTTGTGGAGCGCTAT 376

Db 363 GTGGACTGGCCCTAT 377

RESULT 9  
 BE636646

LOCUS

DEFINITION

rockefeller.0.301 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM320876  
 VERSION  
 EST.

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 482 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..482  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 482;  
 Best Local Similarity 54.8%; Pred. No. 4;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 243 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 302

QY 302 CGCACGTGGATATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 303 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 362

QY 362 TTTGTGGAGCGCTAT 376

Db 363 GTGGACTGGCCCTAT 377

RESULT 9  
 BE636646

LOCUS

DEFINITION

rockefeller.0.301 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM320876  
 VERSION  
 EST.

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 482 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..482  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 482;  
 Best Local Similarity 54.8%; Pred. No. 4;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 243 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 302

QY 302 CGCACGTGGATATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 303 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 362

QY 362 TTTGTGGAGCGCTAT 376

Db 363 GTGGACTGGCCCTAT 377

RESULT 9  
 BE636646

LOCUS

DEFINITION

rockefeller.0.301 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM320876  
 VERSION  
 EST.

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 482 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..482  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 482;  
 Best Local Similarity 54.8%; Pred. No. 4;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 243 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 302

QY 302 CGCACGTGGATATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 303 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 362

QY 362 TTTGTGGAGCGCTAT 376

Db 363 GTGGACTGGCCCTAT 377

RESULT 9  
 BE636646

LOCUS

DEFINITION

rockefeller.0.301 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM320876  
 VERSION  
 EST.

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 482 Std Error: 0.00

POLYA=Yes.

FEATURES  
 source

1..482  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 482;  
 Best Local Similarity 54.8%; Pred. No. 4;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCTGTGAGCCYACCGTGGCGGCGCCAGAGTGGCGCGCCACCGCTCTCTCCGCA 301

Db 243 TCGGTGTATCAAGACCGTCTGCGCAAGGTCCTCCGGCAAGACCGGTGCGCGCGGCA 302

QY 302 CGCACGTGGATATGATGCTGGGCGGCCACCTATGCTCAGCTCTCTACGTAGGAGACC 361

Db 303 AGAAGACGCCAAGAGAGTGAGCAGCGCTCTTGTCTGTGCAACCTCTCTAAGAGAGC 362

QY 362 TTTGTGGAGCGCTAT 376

Db 363 GTGGACTGGCCCTAT 377

RESULT 9  
 BE636646

LOCUS

DEFINITION

rockefeller.0.301 Mastigamoeba balamuthi lambda ZAP II Library  
 Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION  
 VERSION

KEYWORDS  
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM320876  
 VERSION  
 EST.

KEYWORDS  
 SOURCE

ORGANISM  
 Mastigamoeba balamuthi

REFERENCE  
 1 (bases 1 to 482)

AUTHORS  
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
 Philippe,H.

TITLE  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL  
 MEDLINE  
 PUBMED

COMMENT  
 Contact: Muller Miklos  
 Laboratory of Bio

1230 York Avenue, New York, NY 10021, USA  
 Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 571 Std Error: 0.00  
 POLYA=Yes.

# FEATURES

source  
 Location/Qualifiers  
 1..571  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

## ORIGIN

Query Match 8.6%; Score 38.6; DB 2; Length 571;  
 Best Local Similarity 54.8%; Pred. No. 4.1;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;  
 QY 242 TCCCTGTGAGCCYACCGTCGCGGTAAGTCGCGCCGCGCCACCGCTCTCTCCGCA 301  
 Db TCGGTGTCTATCAAGACCGTCGCGCAAGTCCCGGCAAGACCGGTGCGCGCGGCA 377  
 QY 302 CGCAGCTGGATATGATGTGGGCGGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 361  
 Db AGAAGGACGCCCAAGAAGTGAAGTCCCGGCAAGACCGGTGCGCGCGGCA 377  
 QY 362 TTTGTGGAGCGCTAT 376  
 Db GTGGACTGGCCCTAT 452

## RESULT 10

BM321414 650 bp mRNA linear EST 03-JAN-2002  
 LOCUS rockefeller.0.1246 Mastigamoeba balamuthi lambda ZAP II Library  
 DEFINITION Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,  
 mRNA sequence.

ACCESSION BM321414  
 VERSION BM321414.1 GI:18055820  
 KEYWORDS EST.  
 SOURCE Mastigamoeba balamuthi  
 ORGANISM Mastigamoeba balamuthi

REFERENCE 1 (bases 1 to 650)  
 AUTHORS Bapteste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sengen, C.W.,  
 Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and  
 Philippe, H.

TITLE The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba  
 Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

JOURNAL MEDLINE  
 PUBMED  
 COMMENT Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA  
 Email: mmuller@rockvax.rockefeller.edu  
 Insert Length: 650 Std Error: 0.00  
 POLYA=Yes.

## FEATURES

source  
 Location/Qualifiers  
 1..650  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"

## ORIGIN

Query Match 8.6%; Score 38.6; DB 4; Length 650;  
 Best Local Similarity 54.8%; Pred. No. 4.2;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;

QY 242 TCCCTGTGAGCCYACCGTCGCGGTAAGTCGCGCCGCGCCACCGCTCTCTCCGCA 301  
 Db TCGGTGTCTATCAAGACCGTCGCGCAAGTCCCGGCAAGACCGGTGCGCGCGGCA 469  
 QY 302 CGCAGCTGGATATGATGTGGGCGGCCACCCCTATGCTCAGCTCTCTACGTAGGAGAC 361  
 Db AGAAGGACGCCCAAGAAGTGAAGTCCCGGCAAGACCGGTGCGCGCGGCA 529  
 QY 362 TTTGTGGAGCGCTAT 376  
 Db GTGGACTGGCCCTAT 544

## RESULT 11

CL961612 1416 bp DNA linear GSS 21-SEP-2004  
 LOCUS OsIPCC006646 Oryza sativa Express Library Oryza sativa (indica  
 DEFINITION cultivar-group) genomic, genomic survey sequence.

ACCESSION CL961612  
 VERSION CL961612.1 GI:52377967  
 KEYWORDS GSS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)

REFERENCE 1 (bases 1 to 1416)  
 AUTHORS Ma, L., Wang, C., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,  
 Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
 Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and  
 its comparison to Arabidopsis  
 JOURNAL Unpublished (2004)  
 COMMENT Department of Bioinformatic  
 Beijing Institute of Genomics  
 Chinese Academy of Sciences, Beijing 101300, China  
 Tel: 86-10-80481559  
 Fax: 86-10-80488676  
 Email: chenchen@genomics.org.cn  
 Rice genomic sequence.  
 Class: exon-trapped.

## FEATURES

source  
 Location/Qualifiers  
 1..1416  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Express Library"  
 /note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

Query Match 8.6%; Score 38.6; DB 9; Length 1416;  
 Best Local Similarity 53.0%; Pred. No. 4.9;  
 Matches 80; Conservative 1; Mismatches 70; Indels 0; Gaps 0;  
 QY 181 ATCTCCACATTACCTGGCTGTGTCCTCCCTGGTACGCTCTGGCAATACATCAAGATCTGG 240  
 Db ATGTGGACCTGTCTCGCAGGTTGCGCAGCTGCGCGCGCGGCTCTATCAGGGGG 726  
 QY 241 ATCCCTGTGAGCCYACCGTCGCGGTAAGTCGCGCCGCGCGCGCTCTCTCCGC 300  
 Db ATACCGGCGAGGCCCAACTCCGACGTGTGGCGGCGGTCTCGGCGCAGCCAGCTGCGAC 786  
 QY 301 ACGCAGCTGGATATGATGTGGGCGCGGCCA 331  
 Db GGGGAGCTCGACATGGCCCAACGTCGCGCGGA 817

## RESULT 12

BB636716 1538 bp mRNA linear EST 03-JAN-2002  
 LOCUS rockefeller.0.427 Mastigamoeba balamuthi lambda ZAP II Library  
 DEFINITION Mastigamoeba balamuthi cDNA similar to elongation factor 1 alpha,

mRNA sequence.  
 BB636716 1 GI:9919827  
 EST.  
 Mastigamoeba balamuthi  
 SOURCE Mastigamoeba balamuthi  
 ORGANISM  
 Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.  
 1 (bases 1 to 1538)  
 Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Senses, C. W.,  
 Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and  
 Philippe, H.  
 The analysis of 100 genes supports the grouping of three highly  
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba  
 Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)  
 21819461  
 11830664  
 Contact: Muller Miklos  
 Laboratory of Biochemical Parasitology  
 The Rockefeller University  
 1230 York Avenue, New York, NY 10021, USA  
 Email: mmuller@rockefeller.edu  
 Insert Length: 1538 Std Error: 0.00  
 POLYA=Yes.  
 Location/Qualifiers  
 1..1538  
 /organism="Mastigamoeba balamuthi"  
 /mol\_type="mRNA"  
 /strain="ATCC 30984"  
 /db\_xref="taxon:108607"  
 /clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
 /note="syn: Phreatamoeba balamuthi"  
 ORIGIN  
 Query Match 8.6%; Score 38.6; DB 2; Length 1538;  
 Best Local Similarity 54.8%; Pred. No. 4.9;  
 Matches 74; Conservative 1; Mismatches 60; Indels 0; Gaps 0;  
 QY 242 TCCTGTGAGCCYACCGTCGCGTGAAGTCGCCCTGCGCCGCCACCGCTCTCTCCGCA 301  
 Db 1325 TCGGTGTATCAAGACCGTCGTCGCAAGTCCCGCAAGACCGGTGCGCGCGGCA 1384  
 QY 302 CCACGTGTGATATGTTGGGCGGCCACCATATGCTCAGCTCTCTACGTAGAGACC 361  
 Db 1385 AGAAGACGCCAAGAAAGTGAAGCGCGCTCTGCTCTGCAACCTCTCTAAGAGAGC 1444  
 QY 362 TTGTGGAGCGCTAT 376  
 Db 1445 GTGACTGCGCTAT 1459  
 RESULT 13  
 BQ707541/c  
 LOCUS  
 DEFINITION 924 bp mRNA linear EST 16-JUL-2002  
 5', mRNA sequence.  
 ACCESSION BQ707541  
 VERSION BQ707541.1 GI:21846440  
 KEYWORDS  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 924)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail@nih.gov  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LCM2384 row: i column: 12  
 High quality sequence stop: 644.  
 Location/Qualifiers  
 1..924  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6215507"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

# FEATURES source

## ORIGIN

Query Match 8.4%; Score 37.4; DB 5; Length 924;  
 Best Local Similarity 64.4%; Pred. No. 10;  
 Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
 QY 27 TTTACCTGTTGCTCTTCTCTATCTTCCTTCCTGCTTCTCATGCTGCTTACACC 86  
 Db 191 TTTTGTGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 132  
 QY 87 CACACGCCGGCTGGAGTACCGTAATGC 113  
 Db 131 GTCACCCAGGCTGGAGTGCAGTGGTGC 105

## RESULT 14

H29400  
 LOCUS  
 DEFINITION Ym60507.r1 Soares infant brain IN1B Homo sapiens cDNA clone  
 IMAGE:52725 5', mRNA sequence.  
 H29400  
 ACCESSION H29400.1 GI:900310  
 VERSION H29400.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 458)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 1974  
 High quality sequence stops: 319 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1974 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 319.  
 Location/Qualifiers  
 1..458  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:425661"

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source



**THIS PAGE BLANK (USPTO)**